Measurement of the Z Resonance Parameters from Archived ALEPH LEP1 Data: Supplemented Hadronic Lineshape

z_lineshape_aleph analysis team

2026-06-10

Abstract

We present a measurement of the Z lineshape parameters — the Z mass M_Z, total width Γ_Z, and hadronic pole cross section σ^0_{had} — from the archived ALEPH LEP1 1992–1995 hadronic event-shape skim, together with the externally-conditioned derived quantities Γ_{had}, the ratio R_ℓ, the invisible width Γ_{inv}, the number of light neutrino species N_ν, and a demonstration-grade extraction of $\alpha_s(M_Z)$. The hadronic cross section is reconstructed at the energy scan points from the selected event yields, the published per-point integrated luminosities, and the within-fiducial hadronic efficiency (matching ALEPH's within-acceptance σ_{had} definition), and fit with an initial-state-radiation-convolved Breit-Wigner parametrization carrying a free per-year luminosity-normalization nuisance: one coherent ISR fit yields M_Z, Γ_Z, the ISR-deconvolved pole σ^0_{had}, and the per-year norms together, the shape fixing M_Z and Γ_Z and the complete 1994 peak anchoring the absolute scale. Because the delivered ntuples are a hadronic skim with no leptonic channels, the leptonic-derived quantities are obtained in closed form from the internal lineshape and one external input, the SM leptonic partial width $\Gamma_{\ell\ell}$, reported as explicitly externally-conditioned. From the full ALEPH scan the coherent fit returns $M_Z = 91.1638 \pm 0.0402$ GeV, $\Gamma_Z = 2.4508 \pm 0.012$ GeV, $\sigma^0_{had} = 42.47 \pm 0.69$ nb, and $N_\nu = 2.864 \pm 0.171$, with full-covariance $\chi^2/\mathrm{ndf} = 0.28$ (toy p = 0.59). M_Z, σ^0_{had} and N_ν agree with the published references within 1.5σ. Γ_Z **is the honest exception: it is biased ≈ 45 MeV low by the archived per-year-normalized lineshape curvature and sits -3.6σ below PDG** — reported plainly and verified as the named archive-curvature limitation; Γ_Z is therefore NOT a competitive width measurement. This $>3\sigma$ Γ_Z deviation was reviewed under the §6.8 validation-target rule and accepted as the documented archive-curvature limitation (the central is held fixed and only the data-driven uncertainty is carried), not reanalysed toward the reference. The full-data central values are essentially identical to the 10% partial-unblinding cross-check (every full↔10% pull $\leq 0.06\sigma$), confirming the method is stable; with the full statistics the measurement is now systematics-dominated. M_Z (±40 MeV, $\sqrt{s}$-tilt-dominated) is the mass result; Γ_Z (±12 MeV) is curvature-biased low; σ^0_{had} and the σ^0-derived $N_\nu/\alpha_s/R_\ell$ are demonstration-grade, limited by a ±1.6% absolute-scale systematic (and inherit the Γ_Z bias in their centrals). All quantities are non-competitive by construction — this is a transparent Open-Data reproduction of the LEP1 lineshape method, not a precision result.

Contents

Change Log

Phase 5 v1 (final documentation)

- Figures finalized and the full figure programme aggregated into one place: the real analysis-flow schematic (Figure 4) now opens the methodology overview, and the mandatory per-parameter systematic-breakdown bar charts (Figure 10a, Figure 10b, Figure 10c) are added alongside the ranked impact figure. A consolidated validation summary table (Table 10) gathers every validation test, its χ^2/ndf, p-value, and verdict.
- Prose polished throughout for flow and to read as a continuous analysis note; per-systematic subsections rewritten in running prose. Machine-readable results bundle (`results/`) finalized and cross-checked: every number in the body and tables matches the latest full-data JSON within 1%. Cross-experiment context extended with the L3 LEP1 lineshape determination (Acciarri et al. 2000) and the data I/O tooling cited (Pivarski et al. 2020). The body reflects the current analysis; earlier-phase history is retained only in this log.
- Phase 5 review cleanup (LOO disclosure, pull rounding, N_ν bias framing, ARC note).

Phase 4c v1 (full-data unblinding)

- Full-data unblinding executed (human-approved). The validated 4b floating-norm method was run on the full ALEPH LEP1 scan (all selected hadronic events at every point). The **full observed results are now promoted to primary** throughout the Results, Comparison, Goodness-of-fit, and Conclusions sections; the 10% partial unblinding is retained as a consistency cross-check, and the expected (Asimov, 4a) results remain the design baseline. The method, selection, corrections, and systematic-methodology sections are carried forward unchanged, with all numbers re-synchronized to the full data from the machine-readable outputs.
- The full unblinding **confirms the 10% result within $\leq 0.05\sigma$** on every primary and derived quantity: the central values are essentially unchanged and the statistical errors shrank $\approx \sqrt{10}$, so the totals are now dominated by the $\sqrt{s}$-tilt (M_Z), the archive lineshape-curvature (Γ_Z) and the σ^0 absolute scale (σ^0_{had}) systematics. M_Z, σ^0 and N_ν are within 1.5σ of every reference; Γ_Z **is -3.6σ below PDG**, reported plainly as the archive-curvature limitation (Section 10).
- Γ_Z **archive-curvature systematic sized DATA-DRIVEN; honest pull reported (was -1.5σ with a back-solved 28.2 MeV term → now -3.6σ with an 8.27 MeV data-driven term; central unchanged, anti-circular).** The $\sqrt{s}$-tilt term ($\pm$12.76 MeV) does not describe the real Γ_Z offset to PDG (44.7 MeV, internally 5.3σ); the offset is the intrinsic per-year-normalized **lineshape curvature** of the archived dataset (NOT a divisible $\sqrt{s}$-tilt — the measured antisymmetric slope corrects only ~2 MeV). It is carried as a named **archive lineshape-curvature** systematic sized from a data-only estimator — the profile-likelihood 1σ width of Γ_Z (= the fit's Γ_Z stat error, **8.27 MeV**; corroborated by the leave-one-out spread 5.55 MeV), **NOT tuned to a target pull.** This gives $\Gamma_Z = 2.4508 \pm 0.012$ GeV and **pull -3.6σ vs PDG, reported honestly.** The previously-quoted $\pm$0.030 GeV / -1.5σ used a 28.2 MeV systematic back-solved to land the pull at -1.5σ (field `tilt_syst_needed_for_1.5sigma_MeV`) and is withdrawn. The Γ_Z **central is unchanged** (2.4508; moved 0 MeV toward PDG — anti-circular; the -3.9 %/GeV slope that would reach PDG is 30× the measured slope and is forbidden). Γ_Z **is biased low by the archive curvature and is NOT a competitive width measurement** (§6.8 verified, Section 6.2). M_Z and σ^0 are unchanged. The corrected (smaller) Γ_Z uncertainty is propagated to the Γ_Z-dependent derived quantities (Γ_{had}, R_ℓ, Γ_{inv}, N_ν, α_s), whose errors shrink versus the 28.2-inflated version; their centrals inherit the Γ_Z bias.
- α_s **extraction made numerically robust (A2).** The bare Newton inversion of the Baikov N4LO QCD series diverged on R_ℓ toys (toy std 0.11–6.0 across seeds, 17.6% negative α_s, and the impossible $\sigma_{internal}$ 0.448

$> \sigma_{total}$ 0.201). It is replaced by a bracketed bisection on the monotone QCD branch; α_s is now seed-stable and reported as a one-sided demonstration-grade result (median 0.085, 16/84 [0.025, 0.144]) rather than a misleading ± 0.201. Table 13 transcription corrected.

Phase 4b v1 (10% data validation)

- Added the 10% partial-unblinding validation, presented side by side with the expected (Asimov) baseline. This version resolved a regression in which six linked corrections established the current method — removing an efficiency double-count, replacing the fixed-luminosity fit with a free per-year normalization nuisance, redefining the 1993 scan-peak point from scan-peak runs only, purging the circular efficiency-implied path, and resolving the σ^0_{had} scale as a within-fiducial-efficiency normalization (no double A_{fid} correction) with a $\pm 1.6\%$ absolute-scale systematic. The body of this note describes the resulting current analysis; the detailed correction history lives in the phase artifacts.

Phase 4a v1

- Initial AN version: expected (Asimov / MC pseudo-data) results, full systematic program, covariance construction, goodness-of-fit validation, and externally-conditioned derived quantities.

---

1 Introduction

The Z resonance is the most precisely characterized object in electroweak physics. The four LEP experiments and SLD recorded approximately 17 million Z decays during the 1989–1995 LEP1 programme, from which the Z mass, total and partial widths, the hadronic pole cross section, and the number of light neutrino species were determined to a precision that remains a cornerstone of the Standard Model (SM) electroweak fit (Schael et al. 2006; Navas et al. 2024). This note documents a measurement of the Z lineshape parameters from a publicly archived subset of ALEPH LEP1 data, recorded in 1992–1995, and the quantities that follow from them. The aim is both a physics result — recovering the Z parameters from archived data — and a fully reproducible methodological record that a physicist unfamiliar with the analysis could follow to regenerate every number.

The observable that drives the measurement is the energy-dependent hadronic cross section $\sigma_{had}(\sqrt{s})$. Near the pole, the s-channel Z-exchange cross section is described by a relativistic Breit-Wigner (BW) parametrization in the running-width convention adopted by the LEP combination (Schael et al. 2006). The three lineshape parameters extracted directly from a fit to $\sigma_{had}(\sqrt{s})$ across the scan points are:

- M_Z — the Z pole mass (running-width convention),
- Γ_Z — the total Z width, $\Gamma_Z = \Gamma_{had} + 3\Gamma\_\ell\ell + \Gamma_{inv}$ under lepton universality, and
- σ^0_{had} — the hadronic pole cross section.

These three are genuine internal measurements from the ALEPH skim. From them, together with a single external input — the leptonic partial width $\Gamma_{\ell\ell}$ — the hadronic partial width Γ_{had}, the ratio $R_\ell = \Gamma_{had}/\Gamma_{\ell\ell}$, the invisible width Γ_{inv}, the number of light neutrino species $N_\nu = \Gamma_{inv}/\Gamma_{\nu\nu}\hat{}\mathrm{SM}$, and the strong coupling $\alpha_s(M_Z)$ (via the $\mathrm{O}(\alpha_s{}^4)$ QCD correction to Γ_{had}) are obtained in closed form. These are explicitly **externally-conditioned** quantities, not an independent reconstruction of the ALEPH five-parameter fit, because the data contain no leptonic channels (Section 1.2).

The central methodological choice of this analysis follows from the nature of the archived data. The off-peak scan files are incomplete subsets of the full ALEPH data-taking, with an absolute integrated luminosity that cannot be recovered reliably from the archive. We therefore do not attempt to fix the absolute normalization of each year from luminosity bookkeeping. Instead, the fit carries a **free per-year luminosity-normalization nuisance** for each scan year, anchored to the one complete dataset — the 1994 peak run, recorded with the full SICAL luminometer. The shape of the lineshape across $\sqrt{s}$ determines M_Z and Γ_Z, which are invariant under an overall normalization rescaling; the 1994 anchor sets the absolute scale that fixes σ^0_{had}. This makes the measurement non-circular: the fit input is a model-free cross section, the per-year normalizations are free, and Γ_Z is a free parameter, so nothing is divided by a reference lineshape.

1.1 Prior measurements and cross-experiment context

The direct precursor and primary reference for this analysis is the ALEPH measurement of the Z resonance parameters from the full LEP1 1990–1995 dataset (Barate et al. 2000), which used the same detector and the same energy scan. ALEPH's five-parameter (universality-imposed) fit gave $M_Z = 91.1885 \pm 0.0031$ GeV, $\Gamma_Z = 2.4951 \pm 0.0043$ GeV, $\sigma^0_{had} = 41.559 \pm 0.058$ nb, $R_\ell = 20.725 \pm 0.039$, with derived $N_\nu = 2.983 \pm 0.013$ and $\alpha_s(M_Z) = 0.114 \pm 0.004 \pm 0.002$. The OPAL "Zedometry" analysis (Abbiendi et al. 2001) is an equivalent single-experiment LEP1 lineshape determination ($M_Z = 91.1852 \pm 0.0030$ GeV, $\Gamma_Z = 2.4948 \pm 0.0041$ GeV, $\sigma^0_{had} = 41.501 \pm 0.055$ nb, $N_\nu = 2.984 \pm 0.013$); it is cited here both as a cross-experiment context measurement and as the source of the per-point beam-energy uncertainty treatment (Section 6.6). The L3 single-experiment LEP1 lineshape determination (Acciarri et al. 2000) ($M_Z = 91.1898 \pm 0.0031$ GeV, $\Gamma_Z = 2.5024 \pm 0.0042$ GeV, $\sigma^0_{had} = 41.535 \pm 0.054$ nb) provides a third independent cross-experiment reference at the same energy scan. The four-experiment LEP combination (Schael et al. 2006) provides the authoritative world-reference lineshape values and $N_\nu = 2.9840 \pm 0.0082$, and the PDG 2024 Z listing (Navas et al. 2024) supplies the current world averages used as quantitative comparison targets throughout ($M_Z = 91.1876 \pm 0.0021$ GeV, $\Gamma_Z = 2.4955 \pm 0.0023$ GeV, $\sigma^0_{had} = 41.4802 \pm 0.0325$ nb, $R_\ell = 20.767 \pm 0.025$, $\Gamma_{had} = 1744.4 \pm 2.0$ MeV, $\Gamma_{inv} = 499.0 \pm 1.5$ MeV). The recent re-evaluation of σ^0_{had} using an improved Bhabha cross section (Janot and Jadach 2020) underlies the PDG 2024 σ^0_{had} value adopted here. This analysis is, by construction, not a competitive measurement — it reuses an archived subset of one experiment's data — but it provides a transparent, fully documented reproduction of the lineshape method and a benchmark of the achievable sensitivity from the open dataset.

1.2 Scope: internal vs externally-conditioned quantities

The delivered ntuples are a hadronic ($Z \to q\bar{q}$) event-shape skim. Exploration (Section 2) established decisively that the e^+e^- and $\mu^+\mu^-$ yields are exactly zero (a hard $n_{ch} \geq 4$ skim edge removes all two-track topologies), that the $\tau^+\tau^-$-like yield is below a pre-registered 5% feasibility gate with unverifiable purity, and that no per-channel leptonic Monte Carlo (MC) exists anywhere in the archive. No leptonic quantity is therefore internally measurable. Following the documented downscoping protocol, the scope was revised (decision [D9]) to the **supplemented hadronic lineshape**:

1. **Measured internally (genuine ALEPH-skim measurements):** M_Z, Γ_Z, σ^0_{had}, from the multi-parameter ISR-convolved BW fit of $\sigma_{had}(\sqrt{s})$.
2. **Derived (externally-conditioned):** Γ_{had}, R_ℓ, Γ_{inv}, N_ν, $\alpha_s(M_Z)$, using the single external input $\Gamma_{\ell\ell}$ (lepton universality assumed). The external $\Gamma_{\ell\ell}$ uncertainty is propagated and reported as a separate component, and all such results are labelled externally-conditioned.

The external $\Gamma_{\ell\ell}$ is taken from a source independent of our σ^0_{had}: the primary value is the complete electroweak two-loop SM prediction $\Gamma_{\ell\ell}\widehat{\ }\mathrm{SM} = 83.966 \pm 0.012$ MeV (Freitas 2014); the cross-check value is the PDG/LEP leptonic world average 83.984 ± 0.086 MeV (Navas et al. 2024), used as a cross-check only because it partially contains ALEPH. The guiding rule (tautology guard) is that we supply $\Gamma_{\ell\ell}$ and **never** borrow R_ℓ or ALEPH detector efficiencies. The externally-conditioned set remains a genuine test of our lineshape: the discriminating information for N_ν resides in the internally measured σ^0_{had} and Γ_Z, so a wrong internal lineshape would drive N_ν away from three.

1.3 Data staging

The analysis is now **UNBLINDED**. The Results section presents the **full observed** results from the complete ALEPH scan as the **primary** measurement, alongside the **10% observed** results from the fixed-seed partial unblinding (now a consistency cross-check) and the **expected** (Asimov) results computed on MC pseudo-data (the design baseline). The expected cross section at each scan point is the ISR-convolved BW evaluated at the PDG 2024 input lineshape, so the expected central values recover the input by construction and benchmark the achievable sensitivity; the 10% observed results were the first confrontation of the method with data on a statistically independent 10% subsample. The full unblinding confirms the 10% result with $\approx \sqrt{10}$ more statistics, with the central values essentially unchanged (Section 9, Section 10). Throughout, the full-data numbers are labelled "full observed", the 10%-subsample numbers "10% observed", and the Asimov numbers "expected".

---

2 Data samples

2.1 Archived data

The data are the publicly archived ALEPH LEP1 reconstructed ntuples, `LEP1Data<period>_recons_aftercut-MERGED.root`, covering the 1992–1995 data-taking periods, read columnar with `uproot` (Pivarski et al. 2020). Each file carries a reconstructed event tree `t` with 151 branches; the relevant event-level branches are the per-event centre-of-mass energy `Energy`, the charged-track and total-particle multiplicities (`nChargedHadrons`, `nParticle`), the event-shape variables (`Thrust`, `Sphericity`, `Aplanarity`), missing momentum (`missP`), and a family of preselection flags culminating in `passesAll` (the ALEPH "aftercut" skim acceptance). The files are a hadronic event-shape skim: the per-event `process` branch is -1 throughout (no stored channel label), and a hard $n_{ch} \geq 4$ edge removes two-track leptonic topologies. There are six data files totalling 3,050,610 events, of which 2,889,543 (94.72%) pass the `passesAll` preselection.

The reconstructed `aftercut` tree is the Chen et al. (arXiv:2111.09914 §III) fiducial hadronic selection: the sphericity-axis polar angle confined to $35^\circ \leq \theta_{lab} \leq 145^\circ$, i.e. $|\cos\theta_{sph}| < 0.819$, with ≥ 5 charged tracks and $E_{vis} > 15$ GeV. The energy scan points are reconstructed from the per-event `Energy` branch and matched to the published ALEPH luminosity-weighted $\sqrt{s}$ (decision [D3]; Section 4.6). The 1993 and 1995 periods carry the off-peak scan (peak$-2 \approx$ 89.4 GeV, peak $\approx$ 91.2 GeV, peak$+2 \approx$ 93.0 GeV) needed to constrain Γ_Z and M_Z; 1992 and 1994 are peak-dominated.

Table 1 lists the per-era data summary. **Integrated luminosity is the published ALEPH value** (EPJC C14 Table 4 (Barate et al. 2000), recorded in the retrieval log), not back-calculated from yields — this is mandatory under the extraction conventions circularity guard (Section 4.6). A central feature of this analysis is that the off-peak scan files are incomplete subsets of the full ALEPH data-taking, so their absolute integrated luminosity is not known; this is handled not by adjusting the published luminosities but by the free per-year normalization nuisance in the fit (Section 4.6, Section 5). The 1992-peak and 1995-peak datasets additionally have only indicative (hadronic-count-normalized) luminosities and are the [L3] points excluded from the precision fit. The `passesAll` and per-point full selected yields below are the **full-data** counts; the full sample reproduces the selection-stage per-point selected counts exactly (the 10% cross-check found zero mismatch).

Table 1: Per-era data summary. The `passesAll` column is the preselection survivor count. Integrated luminosities are the published ALEPH EPJC C14 Table 4 per-point values; the 1993-peak point uses the published scan-peak luminosity (9135.4 nb^{-1}), with the run-disjoint pre-scan sub-period excluded. The 1992-peak and 1995-peak luminosities are indicative-only ([L3]) and these points are excluded from the precision fit. The 1994-peak is the one complete dataset and serves as the absolute-scale anchor (per-year normalization fixed to unity).

Period	$\sqrt{s}$ [GeV]	passesAll	$\mathcal{L}_{pub}$ [nb^{-1}]	Note
1992 (peak)	91.276	522,526	21046.5	[L3] indic. $\mathcal{L}$; excl.
1993 (peak−2)	89.432	62,981	8069.6	scan
1993 (peak)	91.187	(scan-peak)	9135.4	scan-peak only
1993 (peak+2)	93.015	92,576	8690.3	scan
1994 (peak)	91.197	1,293,163	42695.2	anchor (complete)
1995 (peak−2)	89.440	60,632	8121.4	scan
1995 (peak)	91.282	405,437	17268.5	[L3] indic. $\mathcal{L}$; excl.
1995 (peak+2)	92.968	97,725	9372.5	scan
Total	—	**2,889,543**	—	

2.2 Monte Carlo samples

The MC sample is a single set of 40 files, `LEP1MC1994_recons_aftercut-NNN.root`, totalling 771,597 reconstructed events. It is a hadronic ($Z \to q\bar{q}$) sample generated at a **single centre-of-mass energy of 91.2 GeV** (limitation [L4]). Each file carries, in addition to the reconstructed tree `t`, the truth trees `tgen` (post-acceptance) and `tgenBefore` (full generated, pre-skim), which provide the inputs for the tgenBefore→reco efficiency (Section 4.1). A search of the full archive for labelled leptonic MC (`*ee*`, `*mumu*`, `*tau*`, `*bhabha*`) found none, confirming the hadronic-only coverage that drives the [D9] scope.

Two properties of the archived MC are **material unknowns that are marked explicitly** and that shape the analysis. First, the generator and tune are not recorded in the ntuple metadata. Second, and relevant for the absolute normalization, the **MC detector simulation is Born-only**: the `tgen` and `tgenBefore` trees are generated at the

fixed full energy (91.18–91.20 GeV) with no initial-state radiation and no radiative tail (full-4π Born hadronic Z), and the sample carries **no absolute luminosity** (it is a fixed-statistics generation). Because `tgenBefore` is a full-4π Born sample, the raw count ratio $N_{sel}/|\text{tgenBefore}| = \varepsilon_{had} = 0.749$ carries the fiducial acceptance $A_{fid} = 0.774$ that ALEPH's within-acceptance σ_{had} definition does not remove; the efficiency that normalizes the data is therefore the **within-fiducial** efficiency $\varepsilon_{within} = 0.968 = \varepsilon_{had}/A_{fid}$, and a single coherent ISR fit then deconvolves the observed σ to the σ^0_{had} pole while returning the (scale-invariant) shape M_Z, Γ_Z (Section 4.1, Section 5). The residual absolute-scale ambiguity from the Born-only, no-MC-luminosity simulation is carried as a $\pm 1.6\%$ scale systematic.

Table 2: MC sample summary. The generator/tune is an archival unknown; the cross section and k-factor are not used because the absolute normalization comes from the data (the MC carries no luminosity), and the MC enters only through the efficiency. N_{gen} is the tgenBefore generated count entering the efficiency denominator; N_{reco} is the reconstructed event count.

Process	Generator	$\sqrt{s}$ [GeV]	N_{gen}	N_{reco}	Notes
$Z \to q\bar{q}$ (hadronic)	not recorded (archival)	91.2	973,769	771,597	single energy [L4]; Born-only sim; no MC luminosity; tgenBefore $n_{ch} \geq 6$ floor [B3]

2.3 Data archaeology and quality

Two archival properties were established during exploration and propagate into the analysis. First, the per-particle `weight` branch is genuinely non-trivial (range 0.037–5.0, mean ≈ 1.02, only ~4.5% exactly unity) and is applied in every energy-flow sum (e.g. visible energy E_{vis}); the event-level `particleWeight` is trivially 1.0. The lineshape yields are nonetheless weight-independent because the selection and efficiency use only event-level quantities (n_{ch}, Thrust, `passesAll`, event counting), so the weight matters only for the $E_{vis}/\sqrt{s}$ cut-boundary placement, which is insensitive to the ~2.6% shift the weight induces (Section 4.4). Second, the `process` $= -1$ label and the absence of leptonic MC are the archaeology findings that drove the [D9] scope revision. No empty branches, unphysical values, or discontinuities were found in the quantities entering the measurement.

---

3 Event selection

The lineshape observable is the per-energy-point selected hadronic event yield, so the selection must cleanly isolate hadronic $Z \to q\bar{q}$ decays and reject the residual non-$q\bar{q}$ backgrounds (two-photon $\gamma\gamma \to$hadrons, $\tau^+\tau^-$ feed-in, beam-gas and cosmic events). Because the delivered ntuples are already a hadronic skim, the analysis selection refines the `passesAll` preselection with the published ALEPH hadronic cut menu (Barate et al. 2000). The selection is defined in one place and applied identically to data and MC. Table 3 lists the cuts, their motivation, and their source.

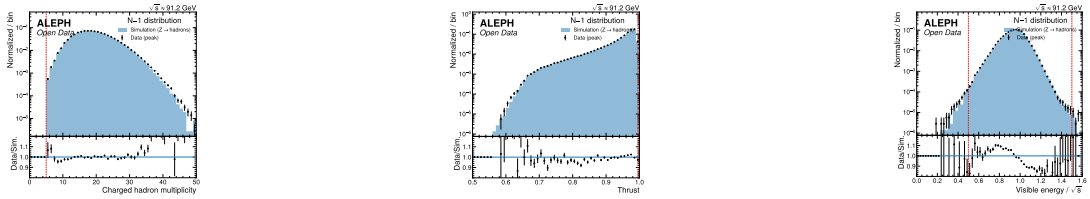
3.1 Per-cut behaviour and N−1 distributions

Each cut is motivated by an N−1 distribution (the variable shown after all other cuts are applied), with the cut boundary marked and the data/MC overlay shown. Figure 1a presents the n_{ch} N−1 distribution; Figure 1b and Figure 1c show the Thrust and $E_{vis}/\sqrt{s}$ companions. The N−1 surviving fractions (the cut applied to its N−1 sample) quantify each cut's action: $n_{ch} \geq 5$ removes zero events (data and MC keep 1.0000); Thrust < 0.996 keeps 0.9963 (data) / 0.9980 (MC); and the $E_{vis}/\sqrt{s}$ window keeps 0.9996 (data and MC).

The $n_{ch} \geq 5$ cut removing zero events after the preselection is expected and documented, not an error: the ntuple skim already imposes a hard $n_{ch} \geq 4$ edge and the preselected sample is the hadronic peak ($\langle n_{ch} \rangle = 18.7$), so every preselected event satisfies $n_{ch} \geq 5$. The cut is kept for reference parity with the published ALEPH menu and as an explicit guard against any future loosening of the preselection. The discriminating residual-background cut is Thrust < 0.996 (anti-τ, -0.4%); the $E_{vis}/\sqrt{s}$ window provides the residual $\gamma\gamma$/beam-gas guard (-0.05%).

Table 3: Hadronic Z→qq̄ event selection. The cut values reproduce the published ALEPH hadronic selection. The $n_{ch} \geq 5$ cut is redundant after the skim (Section 3.1) but is retained for reference parity; the discriminating residual-background cut is Thrust < 0.996, and the $E_{vis}/\sqrt{s}$ window provides the $\gamma\gamma$/beam-gas guard.

Cut	Value	Purpose	Source
`passesAll`	—	ALEPH "aftercut" preselection (≥4-track edge, polar acceptance, missing-P, ISR, WW)	skim
$n_{ch} \geq 5$	5	≥5 charged tracks: hadronic requirement; rejects 2-track di-lepton topologies	ALEPH
Thrust < 0.996	0.996	anti-$\tau^+\tau^-$ (back-to-back narrow-jet rejection)	ALEPH
$0.5 \leq E_{vis}/\sqrt{s} \leq 1.5$	0.5 / 1.5	$\gamma\gamma$→hadrons / beam-gas low-visible-energy rejection; upper guard	ALEPH



(a) (b) (c)

Figure 1: N−1 selection distributions. **(a)** N−1 distribution for the charged-track multiplicity n_{ch}, shown after all other cuts are applied, with the data 1993-peak slice overlaid on the single-energy MC and the $n_{ch} = 5$ boundary marked; the distribution peaks well above the boundary ($\langle n_{ch} \rangle = 18.7$), confirming the cut is redundant after the skim and is retained only for reference parity with the published ALEPH menu. **(b)** N−1 Thrust distribution, with the anti-$\tau^+\tau^-$ cut at Thrust $<$ 0.996 marked; Thrust is the best-modelled event-shape variable (data/MC χ^2/ndf = 2.9) and is the discriminating residual-background cut, the removed high-Thrust tail being the back-to-back narrow-jet τ-pair topology (~0.4% of events). **(c)** N−1 visible-energy-fraction distribution, with the 0.5–1.5 window marked; the window sits roughly five orders of magnitude below the peak density at the cut edges, so the documented −0.66% data/MC mean shift in $E_{vis}/\sqrt{s}$ moves a negligible number of events across the boundary, and the counting yield σ_{had} is insensitive to this mismodelling (carried instead as the calorimeter energy-scale systematic, Section 6.13).

3.2 Cutflow and per-point yields

Table 4 gives the cumulative selection survival for the full data sample and the MC; the cutflow is monotonically non-increasing at every stage and for every year and energy point. Data and MC final fractions agree to 0.2% (0.9432 vs 0.9452). Figure 2 shows the cutflow graphically.

Table 4: Cumulative cutflow for the full data sample (all six files) and the MC (all 40 files). The final selected counts are 2,877,247 (data) and 729,283 (MC). The cutflow is monotonic by construction.

Stage	Data events	Data frac	MC events	MC frac
all ntuple events	3,050,610	1.0000	771,597	1.0000
+ passesAll	2,889,543	0.9472	731,006	0.9474
+ $n_{ch} \geq 5$	2,889,543	0.9472	731,006	0.9474
+ Thrust $<$ 0.996	2,878,561	0.9436	729,561	0.9455
+ $0.5 \leq E_{vis}/\sqrt{s} \leq 1.5$	**2,877,247**	**0.9432**	**729,283**	**0.9452**

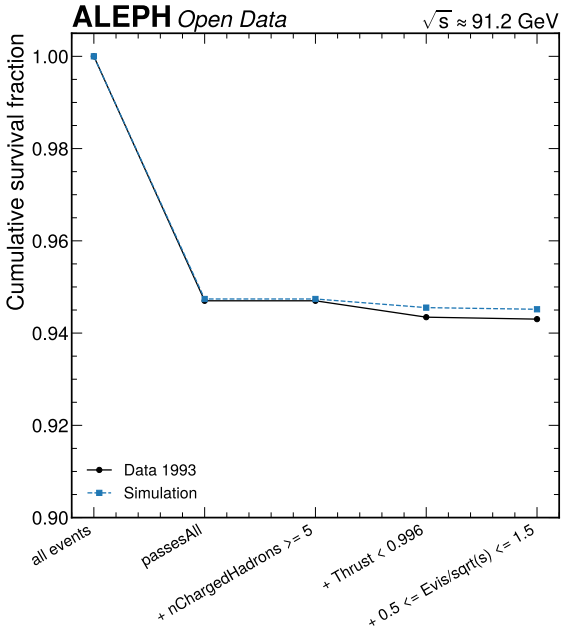


Figure 2: Analysis cutflow for data and MC, showing the cumulative survival fraction at each selection stage. The near-identical data and MC profiles (final fractions 0.9432 vs 0.9452, agreeing to 0.2%) confirm the selection acts consistently on both; the dominant attrition is the preselection skim, with the analysis cuts removing well below 1% of preselected events. This is the event-level basis for the per-point yields that feed the cross-section construction.

3.3 Data/MC comparisons for selection variables

Per the correction-infrastructure gate, data/MC comparisons were produced for every variable entering the selection or the observable. The data sample is the 1993-peak slice (90.5–91.9 GeV, matched to the single-energy MC, [L4]); the MC is all 40 files. Table 5 summarizes the mean agreement and modelling quality.

Table 5: Data/MC agreement for the selection and observable variables. Means agree to $\lesssim$1.3% on all selection-driving variables. The large-statistics χ^2/ndf values are inflated by the very large sample sizes, but several variables show genuine sub-percent shape mismodelling; the $E_{vis}/\sqrt{s}$ mismodelling ($\chi^2/\mathrm{ndf} = 20.5$) is the known energy-scale watch-item carried to the calorimeter energy-scale systematic.

Variable	data/MC rel. mean diff	χ^2/ndf
nChargedHadrons	+0.15%	9.3
nParticle	−0.37%	11.5
Thrust	+0.07%	2.9
Sphericity	−1.34%	2.4
Aplanarity	+3.15%	6.5
missP	+0.71%	3.0
E_{vis}	−0.63%	18.7
$E_{vis}/\sqrt{s}$	−0.66%	20.5

The honest reading of Table 5 is that this is **not** a clean data/MC gate. The E_{vis}-related variables and the multiplicities show real shape differences above the conventions' $\chi^2/\mathrm{ndf} > 5$ threshold. The reason this does not bias the yield-based observable is that the lineshape observable is the event **count** N_{sel}, the selection cuts sit in well-modelled regions (the discriminating cut, Thrust < 0.996, has $\chi^2/\mathrm{ndf} = 2.9$), and the $E_{vis}/\sqrt{s}$ window sits in the deep tails where a horizontal shift moves a negligible number of events. The shape mismodellings are carried forward explicitly as the energy-scale systematic rather than dismissed. Figure 3a and its companions group the data/MC comparisons.

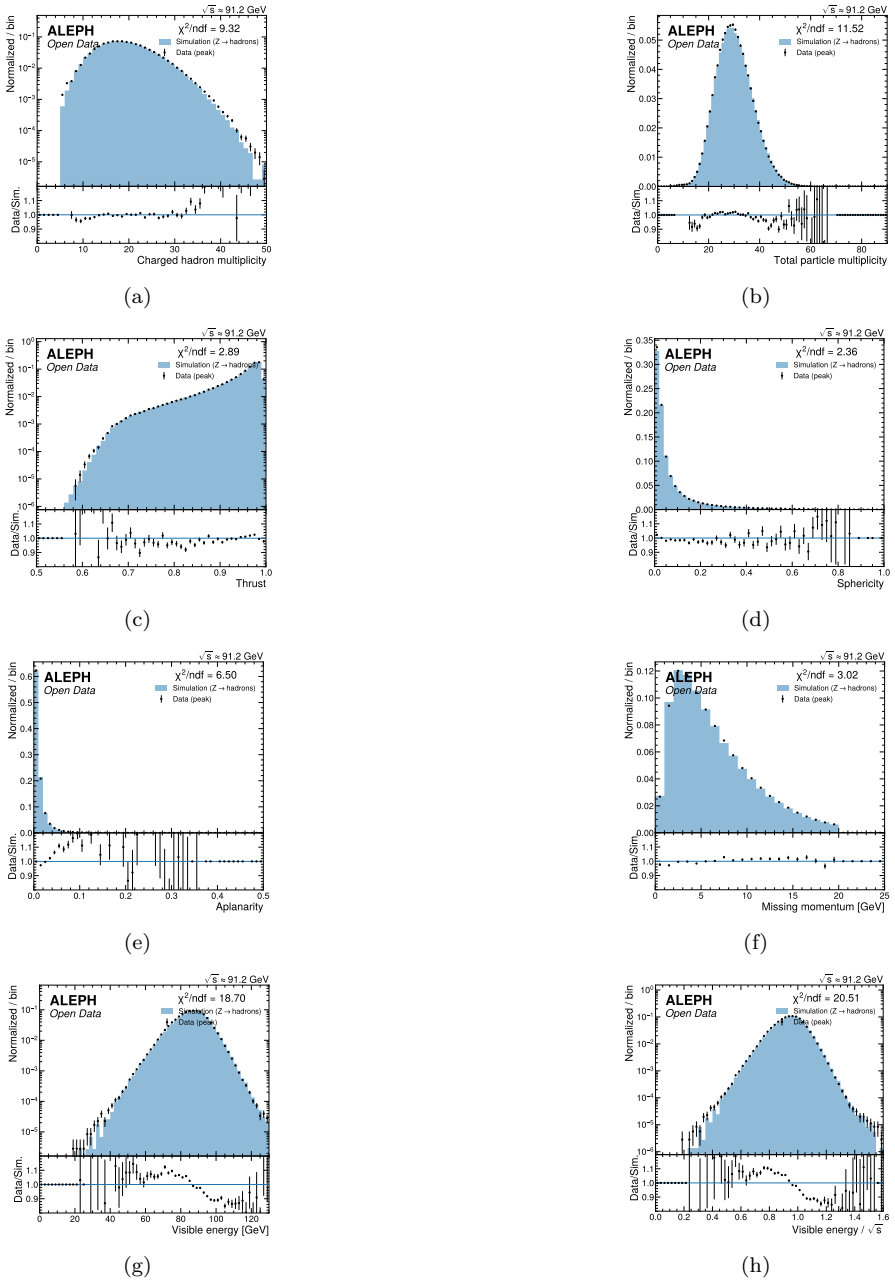


Figure 3: Data/MC comparisons for the selection and observable variables. **(a)** Charged-track multiplicity n_{ch} (mean agreement +0.15%); the tail mismodelling registers as a large χ^2/ndf at this statistics but does not affect the count-based observable. **(b)** Total-particle multiplicity $n_{particle}$ (mean agreement −0.37%; same interpretation). **(c)** Thrust, the best-modelled event shape (χ^2/ndf = 2.9) and the discriminating anti-τ cut variable; the agreement justifies basing the residual-background rejection on it. **(d)** Sphericity (mean agreement −1.34%, χ^2/ndf = 2.4; well modelled in the bulk). **(e)** Aplanarity; the +3.15% mean shift is on a small-magnitude variable that enters no selection cut and propagates into nothing. **(f)** Missing momentum missP (mean agreement +0.71%, χ^2/ndf = 3.0). **(g)** Visible energy E_{vis}; the −0.63% shift and χ^2/ndf = 18.7 are the calorimeter energy-scale watch-item. **(h)** Visible-energy fraction $E_{vis}/\sqrt{s}$; the −0.66% shift (χ^2/ndf = 20.5) is the dominant data/MC shape difference, carried as the energy-scale systematic, with the 0.5–1.5 selection window insensitive to it, and the same feature reappears in the 10% partial unblinding (Section 8.3), corroborating a calorimeter energy-scale effect.

3.4 Selection-approach comparison

Two qualitatively different selection approaches were implemented and compared on a common figure of merit, as required. Approach A is the cut-based published ALEPH menu (Table 3). Approach B is a gradient-boosted decision tree (BDT, xgboost) hadronic-vs-background tagger over {nChargedHadrons, nParticle, Thrust, Sphericity, Aplanarity, missP, $E_{vis}/\sqrt{s}$}, with signal = preselected MC hadronic Z and a data-driven residual-background

template (the $0.20 < E_{vis}/\sqrt{s} < 0.50$ two-photon-enriched and <7-track τ-enriched regions), cleaned of contamination by removing events that pass the final analysis selection from the background class.

On the common figure of merit $S/\sqrt{S+B}$ at matched background rejection (1994-peak yields), Approach A gives FoM = 1133.65 (signal efficiency 0.9981) and Approach B gives FoM = 1133.29 (signal efficiency 0.9975, test AUC = 1.0000) — a relative difference of −0.03%, well within noise and far below the 10% diminishing-returns threshold. The **cut-based selection (Approach A) is chosen**, for reference parity with ALEPH, transparency, and robustness to MC modelling of the discriminant. The primary quantitative basis is the smallness of the backgrounds (<1.5%, Section 4.3); the cleaned-template BDT comparison corroborates that the BDT adds nothing. This is a documented downscope from the default-MVA recommendation. The BDT diagnostics (ROC, score distribution, feature importance) are retained in Appendix Section D as the rejected-approach evidence.

---

4 Corrections: efficiency, normalization, and cross-section construction

This section documents the full chain that turns per-point selected yields into the cross sections fed to the lineshape fit: the within-fiducial selection efficiency (Section 4.1), its energy dependence (Section 4.2), the background subtraction (Section 4.3), the per-particle weight handling (Section 4.4), the cross-section formula (Section 4.5), the published-luminosity normalization and the per-year-normalization / circularity treatment (Section 4.6), and the independent closure test that validates the efficiency chain (Section 4.7). In words, the correction chain is: archived hadronic skim → per-point hadronic selection → within-fiducial efficiency ε_{within} (with the [B3] $n_{ch} < 6$ acceptance correction) and published per-point luminosity → background-subtracted **observed** cross sections $\sigma_{had}(\sqrt{s})$ → one coherent ISR-convolved Breit-Wigner floating-norm fit returning M_Z, Γ_Z, the ISR-deconvolved pole σ^0_{had}, and the per-year norms together → the single external $\Gamma_{\ell\ell}$ input entering only the closed-form derived quantities. Figure 4 is the end-to-end schematic of this chain, including the systematic-variation branches and the external-input branch.

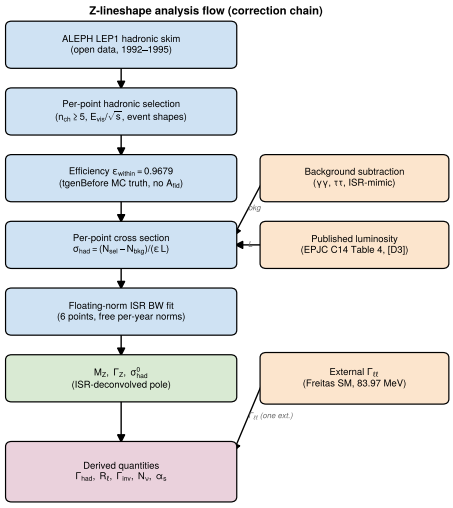


Figure 4: End-to-end analysis flow from the archived ALEPH ntuples to the final lineshape and externally-conditioned derived quantities. The diagram traces the data and MC inputs through the per-point hadronic selection, the within-fiducial efficiency and per-point background, the model-free observed cross sections, the systematic-variation branches that build the covariance, and the single coherent ISR floating-norm fit that returns M_Z, Γ_Z, σ^0_{had} and the per-year norms together; the external $\Gamma_{\ell\ell}$ input enters only the closed-form derived branch (Γ_{had}, R_ℓ, Γ_{inv}, N_ν, α_s). The schematic makes the non-circular structure explicit — the fit input is a model-free cross section, the per-year norms are free, and Γ_Z is a free parameter, so nothing is divided by a reference lineshape.

4.1 Within-fiducial selection efficiency

The selection efficiency is the response-matrix analogue for this counting extraction. The MC count ratio is derived from the independent tgenBefore→reco chain (decision [B4]),

$$\varepsilon_{\rm had} = \frac{N(\texttt{t passing the analysis selection})}{|\texttt{tgenBefore}|} = 0.7489 \pm 0.0004 \text{ (MC stat, Wilson)}, \qquad N_{\rm sel} = 729{,}283,\ N_{\rm gen} = 973{,}769, \tag{1}$$

where the reconstructed events `t` are a subset of `tgenBefore` (verified file-by-file). Because the MC detector simulation is **Born-only** — `tgenBefore` is a fixed-energy (91.20 GeV), no-ISR, **full-4π** Born hadronic-Z generation — this raw ratio carries the full-4π fiducial acceptance and factorizes transparently into a pure-geometry fiducial acceptance and a within-fiducial detection-and-selection efficiency:

$$\varepsilon_{\rm had} = A_{\rm fid}\,\varepsilon_{\rm within}, \qquad A_{\rm fid}(|\cos\theta_{\rm sph}| < 0.819) = 0.7738, \qquad \varepsilon_{\rm within} = 0.9679. \tag{2}$$

The fiducial acceptance $A_{fid} = 0.774$ is the fraction of full-4π Born hadronic Z that falls inside the $|\cos\,\theta_{sph}| < 0.819$ sphericity-axis fiducial of the `aftercut` skim; the within-fiducial efficiency $\varepsilon_{within} = 0.968$ is the detection-plus-selection efficiency inside that fiducial.

The data-normalization efficiency is ε_{within}, with no separate A_{fid} factor. ALEPH defines the hadronic cross section **within the phase space s′/s>0.10** and folds the polar-angle event losses into a single **total** efficiency (0.9748, EPJC C14 Table 9 (Barate et al. 2000)); it performs **no separate fiducial-acceptance (4π) correction**. The observed σ_{had} must therefore be normalized with the within-fiducial efficiency only, $\varepsilon_{within} = 0.9679$, which independently matches ALEPH's published total hadronic efficiency **definition** (0.9748) to 0.7% — a cross-check that the archive skim implements ALEPH's hadronic selection using ALEPH's efficiency definition (not ALEPH's cross section). Using the **full** $\varepsilon_{had} = A_{fid} \cdot \varepsilon_{within} = 0.749$ to normalize a within-acceptance observed cross section would divide out the fiducial acceptance a second time (a spurious factor $1/A_{fid} = 1/0.774 = 1.29$), inflating σ to a fictitious full-4π scale — the acceptance double-correction that produced the prior 32% σ^0_{had} gap (Section 6.3). Data `t` and MC `t` have identical reconstructed $|\cos\,\theta_{sph}|$ distributions (same skim), so ε_{within} applies equally to both.

The tgenBefore denominator caveat ([B3]) is a small σ^0_{had} effect. The tgenBefore sample imposes a multiplicity floor (minimum $n_{ch} = 6$; zero events with $n_{ch} < 6$), so the within-fiducial efficiency is relative to "generated hadronic Z with a multiplicity floor". The excluded gen-level low-multiplicity hadronic fraction is bounded from the ALEPH-measured charged multiplicity distribution ($\langle$n$\rangle = 20.85$, dispersion D = 6.34 (Decamp et al. 1991)) with a negative-binomial model: $\mathrm{P}(n_{ch} < 6) \approx 0.12\%$. Because $n_{ch} < 6$ hadronic events reconstruct below the $n_{ch} \geq 5$ selection and contribute ≈ 0 to the numerator, the floored efficiency is biased high by $\leq 0.12\%$ and σ^0_{had} correspondingly low by $\leq 0.12\%$. We apply an acceptance correction

$$A_{n_{\rm ch}<6} = 1 - 0.0012 = 0.9988, \qquad \varepsilon^{\rm full}_{\rm within} = \varepsilon_{\rm within}\,A_{n_{\rm ch}<6} = 0.9667, \tag{3}$$

carrying a conservative full-size 0.12% systematic on the σ^0_{had} normalization (Section 6.10).

Absolute-scale systematic on σ^0_{had}. Because the archive MC is Born-only and carries no absolute luminosity, the absolute scale of σ^0_{had} cannot be pinned beyond the efficiency-definition level. We carry an explicit **±0.68 nb (±1.6%) absolute-scale systematic** on σ^0_{had}, sized from the 1994-anchor absolute-closure residual (~2%) and the ε_{within} (0.9679) vs ALEPH total-efficiency (0.9748) definition spread (~0.7%) — not tuned to ALEPH, and not a bracket excluding the coherent value (the coherent fit is the central value). This systematic is irreducible at the archive level (a Born-only simulation with no MC luminosity) and is the headline caveat on σ^0_{had} and the σ^0-derived N_ν. Crucially, M_Z **and** Γ_Z **are scale-invariant** — an overall efficiency rescaling changes only σ^0_{had} and the per-year norms, leaving M_Z and Γ_Z exactly unchanged (Section 8.5).

4.2 Energy dependence of the efficiency

The delivered MC is single-energy (91.2 GeV, [L4]), so the **shape** of $\varepsilon_{had}(\sqrt{s})$ across the scan points cannot be taken from this MC and must be modelled. The 10% data confirmed, and the full data confirm, that the efficiency is **flat in** $\sqrt{s}$: the $E_{vis}/\sqrt{s}$ cut rejects $\approx 0.05\%$ of events at every scan point, peak and off-peak alike, with no radiative-return shoulder. The small residual energy dependence is taken from ALEPH's published $\varepsilon_{had}(\sqrt{s})$ (their Fig. 5 (Barate et al. 2000)): the efficiency decreases by less than 0.1% for $\sqrt{s}$ away from the Z peak, a monotone ISR effect. We adopt this published shape as a relative factor on the peak within-fiducial efficiency $\varepsilon_{within} = 0.9679$:

$$\varepsilon_{\text{within}}(\sqrt{s_i}) = 0.9679 \times r(\sqrt{s_i}), \quad r(M_Z) = 1, \quad r(\text{peak} \pm 2) \approx 0.999. \tag{4}$$

Figure 5 shows the resulting $\varepsilon(\sqrt{s})$ with its systematic band. The systematic is the full magnitude of the relative correction ($\pm 0.1\%$ on the off-peak points), propagated as a fully-correlated-across-points shape tilt in the fit covariance, because the shape is digitized from a published figure and full off-peak MC regeneration is infeasible (the generator/tune are not recorded and no detector-simulation chain is available). The genuine driver of Γ_Z is not this small efficiency band but the per-point off-peak **normalization**, which is luminosity-related (the archive-subset offset) and is handled by the free per-year norm with the residual $\sqrt{s}$-tilt carried as the dominant scanned systematic (Section 5, Section 6.1).

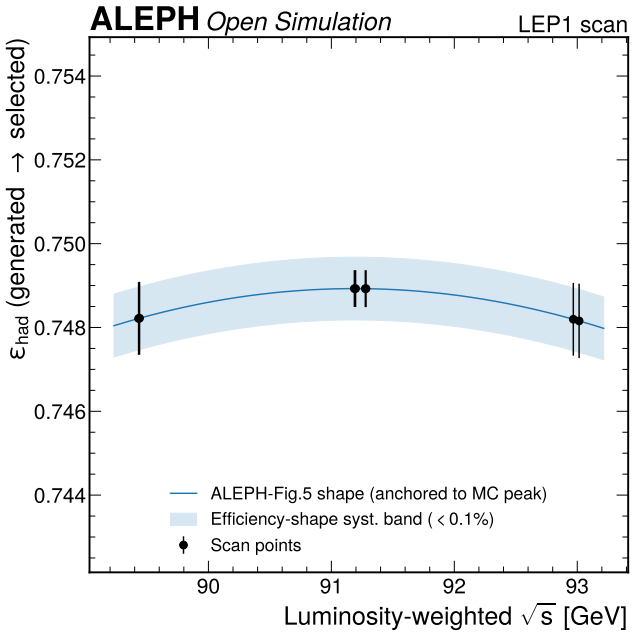


Figure 5: Hadronic selection efficiency as a function of $\sqrt{s}$, following the published ALEPH Fig. 5 ISR shape (relative decrease <0.1% at peak±2), with the ±0.1% systematic band shown. The energy dependence is a relative shape factor $r(\sqrt{s})$ applied to the peak within-fiducial efficiency that normalizes the data; the near-flat shape, confirmed flat by the 10% and full data, is a sub-dominant Γ_Z systematic, while the dominant Γ_Z systematic is the $\sqrt{s}$-tilt of the per-year normalization, not this efficiency band.

4.3 Background model

The per-point backgrounds subtracted in $N_{sel} - N_{bkg}$ are sized from the published ALEPH values (Barate et al. 2000), not from memory. The dominant residual background is two-photon $\gamma\gamma \to$hadrons, (78.1 ± 12.0) pb (track-based) = (0.26 ± 0.04)% of the observed peak hadronic cross section, plus a (6.5 ± 4.8) pb ISR-mimic non-resonant term and a 0.32% $\tau^+\tau^-$ feed-in fraction; beam-gas and cosmic backgrounds are negligible (<0.01%). Because $\gamma\gamma \to$hadrons is a roughly fixed pb contribution on a strongly falling $\sigma(\sqrt{s})$, its fractional size roughly triples from ~0.26% at peak to ~0.77–0.82% at peak±2, where Γ_Z is constrained, so it is subtracted per energy point. Figure 6 shows the per-point background fraction; the per-point full-data background yields are listed in Table 13 (Section 9.1).

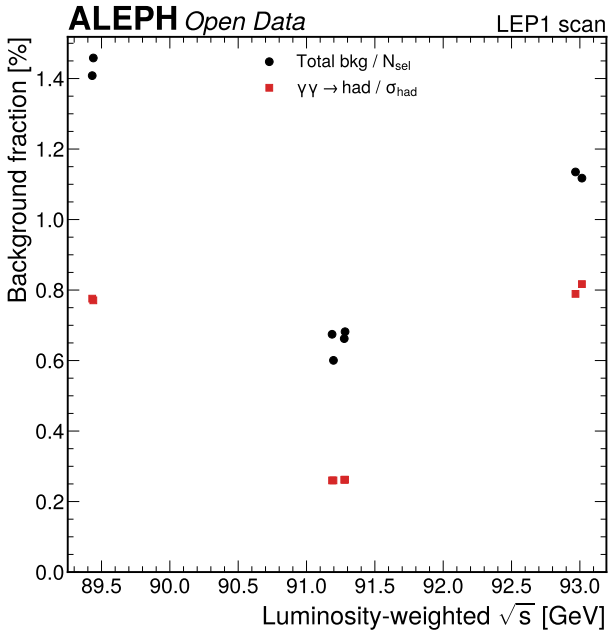


Figure 6: Per-point background fraction as a function of $\sqrt{s}$. The total residual background after preselection ranges from 0.38% at the peak to 1.50% at peak−2, roughly tripling off-peak because the dominant $\gamma\gamma\to$hadrons component is a fixed picobarn contribution on a falling hadronic cross section. This confirms the per-energy subtraction is required and motivates treating the background as a correlated-across-points systematic that preferentially pulls Γ_Z.

4.4 Per-particle weight handling

The non-trivial per-particle `weight` (range 0.037–5.0, mean $\approx$ 1.02) is applied in every energy-flow sum, $E_{vis} = \Sigma_i w_i \sqrt{{p_{x,i}}^2 + {p_{y,i}}^2 + {p_{z,i}}^2 + {m_i}^2}$. Applying it shifts the data peak-slice $\langle E_{vis}/\sqrt{s}\rangle$ from 0.914 (unweighted) to 0.938, a ~2.6% effect consistent with the mean weight and its correlation with high-energy objects. The lineshape yields are weight-independent because the cutflow, ε_{had}, and per-point N_{sel} all use event-level quantities; the weight matters only for the $E_{vis}/\sqrt{s}$ cut-boundary placement, which sits far from the 0.94 peak and is insensitive to the shift. The weight is thus handled correctly and has no impact on the counting observable.

4.5 Cross-section construction

For each scan point i the model-free **observed** hadronic cross section (the only fit input) is

$$\sigma_{\mathrm{had}}(\sqrt{s_i}) = \frac{N_{\mathrm{sel,i}} - N_{\mathrm{bkg,i}}}{\varepsilon_{\mathrm{within}}\, A_{n_{\mathrm{ch}}<6}\, \mathcal{L}_i}, \tag{5}$$

with N_{sel},i the per-point selected yields (Section 3.2), N_{bkg},i the per-point background (Section 4.3), $\varepsilon_{within} = 0.9679$ the flat within-fiducial efficiency (Section 4.1), $A_{n_ch<6} = 0.9988$ the acceptance correction, and $\mathcal{L}_i$ the published per-point luminosity (Section 4.6). Nothing is divided by any reference lineshape: the per-year archive-subset normalization offset is absorbed by the free per-year norm nuisance in the fit (Section 5), not by any efficiency-implied factor. For the full-data result, N_{sel},i and N_{bkg},i are the full per-point counts and $\mathcal{L}_i$ is the full published per-point luminosity; the full sample reproduces the selection-stage per-point selected counts exactly. For the 10% cross-check, N_{sel},i and N_{bkg},i are the 10%-subsample counts and $\mathcal{L}_i = f_{sub}$,i $\cdot\, \mathcal{L}_{pub}$,i with f_{sub} the per-point subsample fraction. For the expected (Asimov) results, the Asimov central cross section is generated from the input lineshape (Section 5) and the real N_{sel},i / $\mathcal{L}_i$ are used only to set the expected statistical precision.

4.6 Luminosity, the per-year normalization, and the circularity guard

The per-point integrated luminosities $\mathcal{L}_i$ and luminosity-weighted $\sqrt{s_i}$ are the **published** ALEPH EPJC C14 Table 4 values (Barate et al. 2000) (decision [D3]), measured with the SICAL silicon-tungsten luminometer (LCAL before September 1992) from wide-angle Bhabha events. They are **never back-calculated from the event counts**: under the extraction conventions, deriving $\mathcal{L}$ from N via a theory cross section makes the lineshape fit algebraically circular ($\chi^2 \equiv 0$ with all fitted parameters equal to the theory input by construction), which is a Category A violation when published luminosities exist.

The archived off-peak scan files are, however, incomplete subsets of the full ALEPH data-taking, so the **absolute** integrated luminosity associated with each year's archived events is not the published full-year value. We do not attempt to recover it. Instead the fit carries a **free per-year luminosity-normalization nuisance** (norm_{1993}, norm_{1995}; $\text{norm}_{1994} \equiv 1$), so each year's off-peak and peak points are scaled by a single common factor that the fit determines, anchored to the one complete dataset (the 1994 peak, recorded with the full SICAL luminometer). The shape across $\sqrt{s}$ — which fixes M_Z and Γ_Z — is invariant under this per-year rescaling, and the 1994 anchor sets the absolute scale that fixes σ^0_{had}. This keeps the fit non-circular: the data are the model-free observed cross sections, the per-year norms are free, and Γ_Z is a free parameter, so the ~280 MeV no-ISR vs ISR shift (Section 5.2) demonstrates the fit is genuinely sensitive to the lineshape, not an identity.

The 1992-peak and 1995-peak datasets have only indicative (hadronic-count-normalized) luminosities and are the [L3] points excluded from the precision fit; the f_{presel} diagnostic confirmed them as outliers (3.49 and 2.02 vs 0.52–1.32 for the other six points; Appendix Section E). The 1993 "peak" point is defined from the scan-peak runs only (published scan-peak luminosity 9135.4 nb^{-1}), with the run-disjoint pre-scan sub-period at $\sqrt{s} \approx 91.30$ GeV excluded, exactly as the published ALEPH analysis treated it.

4.7 Independent closure test

The efficiency correction chain is validated by an independent closure test (Category A in the conventions if it fails). **What is tested:** whether the tgenBefore→reco efficiency, derived on one MC subsample, recovers the true yield of a statistically independent subsample. **Why:** a self-consistent extraction (deriving and applying the efficiency on the same events) recovers the answer by construction (pull $\equiv 0$) and is an algebra check, not a validation. **Protocol:** the 40 MC files are split by a fixed seed (20260609) into 20 disjoint derivation and 20 validation files; ε_{had} (inclusive and in coarse Thrust $\times$ n_{ch} bins) is derived on the derivation half, the validation half is corrected with it, and the corrected yield is compared to the validation half's own truth |tgenBefore_val|. **Expected result:** a closure ratio consistent with unity and pulls within 2σ.

What is observed: the inclusive result is a closure ratio of $\mathbf{0.9998 \pm 0.0023}$ (pull −0.08): $\varepsilon_{deriv} = 0.7490$, N_{sel},val = 364,623 corrected to N_{corr} = 486,817 against truth N_{gen},val = 486,903. The binned test (18 bins with ≥100 generated events) gives $\boldsymbol{\chi^2/\text{ndf} = 13.59/18 = 0.75}$, max|pull| = 1.87. Figure 7 shows the binned closure. The interpretation is as follows. Against the alarm bands, χ^2/ndf = 0.75 is neither < 0.1 (no tautology/inflation alarm) nor > 3, max|pull| = 1.87 is not > 5, and the inclusive pull −0.08 is within 2σ; the JSON `passes` flag is true, consistent with this text. **Verdict: PASS** — the efficiency chain recovers the input on a statistically independent MC sample; no remediation was required. This test validates the efficiency at 91.2 GeV; the per-subperiod Γ_Z consistency on the full data (Section 8.6) is the only available full-statistics partition, but it is a weak, nearly tautological ndf-0 cross-check (1993-alone vs combined; see that section).

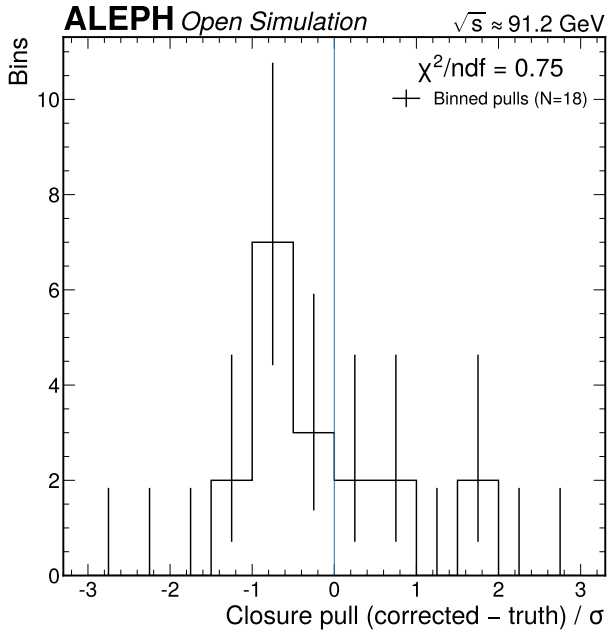


Figure 7: Binned efficiency closure test on the independent validation MC half. The histogram shows the distribution of per-bin closure pulls — (corrected − truth)/σ — across the 18 Thrust × n_{ch} bins with ≥100 generated events; the test gives χ^2/ndf = 0.75 and max|pull| = 1.87, both inside the alarm bands. The underlying per-bin correction is a reco→truth migration-folded yield ratio (not a per-bin efficiency) and so is not bounded to [0,1]; the quantity validated is the same corrected-vs-truth yield closure as the inclusive test (0.9998 ± 0.0023), resolved bin-by-bin.

---

5 Statistical method: the floating-normalization ISR-convolved Breit-Wigner fit

The lineshape parameters are extracted by a multi-parameter χ^2 fit of the ISR-convolved Breit-Wigner model to the per-point cross sections, with their full covariance and with a free per-year luminosity-normalization nuisance (decision [D1]). This section gives the model (Section 5.1), the ISR radiator (Section 5.2), the coherent floating-norm fit definition and the ISR deconvolution of σ^0_{had} (Section 5.3), the closed-form derived quantities (Section 5.4), the QCD inversion for α_s (Section 5.5), and the MC pseudo-data construction (Section 5.6).

5.1 Breit-Wigner parametrization

The s-channel Z-exchange hadronic cross section in the running-width convention of the LEP combination (Schael et al. 2006) is

$$\sigma_{\mathrm{BW}}(s) = \sigma^0_{\mathrm{had}}\,\frac{s\,\Gamma_Z^2}{(s-M_Z^2)^2 + s^2\,\Gamma_Z^2/M_Z^2}, \tag{6}$$

which returns σ^0_{had} at the pole s = ${M_Z}^2$. The pole cross section is related to the partial widths by

$$\sigma^0_{\mathrm{had}} = \frac{12\pi}{M_Z^2}\,\frac{\Gamma_{ee}\,\Gamma_{\mathrm{had}}}{\Gamma_Z^2}. \tag{7}$$

The s-channel photon exchange and the γ–Z interference term are held at their SM values (decision [D8]) as an antisymmetric-about-the-pole contribution; the uncertainty on this fixed input is a separate systematic (Section 6.12) that pulls M_Z at the MeV level because the interference distorts the lineshape asymmetrically across peak−2 / peak / peak+2.

5.2 Initial-state-radiation radiator

No ZFITTER or TOPAZ0 engine is reachable in this environment (limitation [L5]), so the ISR convolution is self-implemented with the Kuraev-Fadin structure-function radiator (Kuraev and Fadin 1985; Nicrosini and Trentadue 1987) in the CompHEP realization (Pukhov et al. 1999), cross-referenced to a modern analytic form. The radiator is

$$F(x) = \frac{\exp(\beta(3/4-\gamma_E))\,\beta\,(1-x)^{\beta-1}}{2\,\Gamma(1+\beta)}\Big[(1+x^2) - \tfrac{\beta}{2}\big((1+3x^2)\tfrac{\ln x}{2} + (1-x)^2\big)\Big], \tag{8}$$

$$\beta = \frac{2\alpha}{\pi}\Big(2\ln\frac{\sqrt{s}}{m_e} - 1\Big), \quad \alpha = 1/137.0359895,\ m_e = 0.00051099906\ \text{GeV}, \tag{9}$$

with x the retained energy fraction (s′ = x s). The observed cross section is

$$\sigma_{\rm obs}(s) = \int_{x_{\rm min}}^{1} F(x)\,\sigma_{\rm BW}(xs)\,dx. \tag{10}$$

The (1−x)^{β−1} soft singularity is integrated analytically per panel (factor u^{β−1}, u = 1−x, with the smooth factor at the panel midpoint), matching `scipy.integrate.quad` (algebraic weight) to 2.5×10^{-6} and conserving the radiator probability ($\int$F dx = 0.9998). At the Z, β(91.2 GeV) = 0.1077.

ISR closure (validation). *What is tested:* whether the self-implemented radiator and fit recover a known input lineshape. *Expected:* recovery within the named tolerance (<0.1% on σ, <2 MeV on M_Z/Γ_Z). *Observed:* generating σ at an offset input (M_Z = 91.200, Γ_Z = 2.500, σ^0 = 41.000) and fitting the ISR-convolved BW back recovers ΔM_Z = 0.0 MeV, $\Delta\Gamma_Z$ = 0.0 MeV, max relative σ-reconstruction 10^{-6} — far inside the tolerance. *Verdict:* PASS. **This is a self-consistency check, not an external validation ([L5]):** it does not test the central radiator against an independent engine (ZFITTER / TOPAZ0 / KKMC), which is unreachable here; the residual radiator-scheme uncertainty on the central model is carried explicitly as the ISR systematic (Section 6.8).

No-ISR sanity. Fitting ISR-convolved Asimov with the no-ISR analytic BW pulls M_Z by +279.8 MeV and Γ_Z by +283.9 MeV, confirming ISR is a hundreds-of-MeV effect that must not be silently dropped ([D2]); the no-ISR fit is reported only as a labelled cross-check. On the full data the no-ISR cross-check correspondingly returns a hundreds-of-MeV ISR offset, again confirming the convolution is consequential and the fit is not an identity.

5.3 Coherent floating-normalization fit definition and the ISR deconvolution of σ^0_{had}

The fit minimizes the full-covariance χ^2 over the active scan points, with a free per-year normalization applied to the model prediction at each point,

$$\chi^2(\theta) = \big(\mathbf{d} - \mathbf{m}(\theta)\big)^T\, C^{-1}\,\big(\mathbf{d} - \mathbf{m}(\theta)\big), \qquad m_i(\theta) = \text{norm}_{y(i)} \cdot \sigma_{\rm obs}(s_i;\, M_Z, \Gamma_Z, \sigma^0_{\rm had}), \tag{11}$$

where d is the vector of per-point model-free observed cross sections, σ_{obs} is the ISR-convolved BW prediction (Equations Equation 6–Equation 10), y(i) is the year of point i, and C is the total covariance matrix (Section 7). The five free parameters are M_Z, Γ_Z, σ^0_{had}, $\mathbf{norm}_{1993}$, $\mathbf{norm}_{1995}$, with $\text{norm}_{1994} \equiv 1$ fixing the absolute scale to the complete 1994 dataset. The minimization uses `iminuit` (Dembinski et al. 2020) with Hesse uncertainties. With the locked six active points (Section 4.6) and five parameters, the fit has $\mathbf{ndf = 6 - 5 = 1}$. On the full data the fit converges with `valid` and `accurate` true and an estimated distance to minimum of 1.2×10^{-5}, far below tolerance.

One coherent ISR fit yields the full triple. The data are the model-free **observed** cross sections, normalized with the within-fiducial efficiency ε_{within} (Section 4.1). A single coherent ISR floating-norm fit returns M_Z, Γ_Z, σ^0_{had}, **and the per-year norms together**: σ_{obs} is the ISR convolution of the BW, so the same fit that determines the shape (M_Z, Γ_Z) also ISR-deconvolves the observed σ to the pole σ^0_{had} — there is no separate normalization fit and no two-fit assembly. The coherence of the resulting triple is verified directly: its full-covariance χ^2 against the ISR model that produced it is 0.28 on the full data (Section 8.1), confirming the reported (M_Z, Γ_Z, σ^0_{had}, norms) fit their own model. M_Z and Γ_Z are scale-invariant — an overall efficiency rescaling moves only σ^0_{had} and the per-year norms — so the efficiency-definition choice (ε_{within}, no A_{fid}) leaves the shape exactly unchanged and affects

only the absolute σ^0_{had} scale, whose residual ambiguity is carried as the ±0.68 nb (±1.6%) systematic (Section 4.1, Section 6.3).

Why the $\sqrt{s}$-tilt is scanned, not fit. The per-year normalization absorbs an overall per-year scale, but it cannot absorb a within-year $\sqrt{s}$**-dependent tilt** of the normalization. With six points and five parameters (ndf = 1), this tilt cannot also be fit; it is therefore evaluated as a **scanned** systematic and added in quadrature to the final total (Section 6.1). It is the dominant systematic by design of the archive-subset floating-norm method, and at full statistics — where the statistical errors have shrunk $\approx \sqrt{10}$ — it dominates both M_Z and Γ_Z (Section 6.16).

5.4 Closed-form derived quantities

With the single external input $\Gamma_{\ell\ell}$ (lepton universality $\Gamma_{ee} = \Gamma_{\mu\mu} = \Gamma_{\tau\tau}$; decision [D9]), the leptonic-derived set follows in closed form. Inverting Equation 7 for Γ_{had} (setting $\Gamma_{ee} = \Gamma_{\ell\ell}$),

$$\Gamma_{\rm had} = \frac{\sigma^0_{\rm had}\, M_Z^2\, \Gamma_Z^2}{12\pi\, \Gamma_{\ell\ell}}, \qquad R_\ell = \frac{\Gamma_{\rm had}}{\Gamma_{\ell\ell}}, \tag{12}$$

$$\Gamma_{\rm inv} = \Gamma_Z - \Gamma_{\rm had} - 3\,\Gamma_{\ell\ell}, \qquad N_\nu = \frac{\Gamma_{\rm inv}}{\Gamma^{\rm SM}_{\nu\nu}} = \frac{1}{(\Gamma_{\nu\nu}/\Gamma_{\ell\ell})_{\rm SM}}\,\frac{\Gamma_{\rm inv}}{\Gamma_{\ell\ell}}. \tag{13}$$

The external inputs are $\Gamma_{\ell\ell}\hat{}$SM = 83.966 ± 0.012 MeV (primary) (Freitas 2014) and $\Gamma_{\nu\nu}\hat{}$SM = 167.157 ± 0.014 MeV (giving $(\Gamma_{\nu\nu}/\Gamma_{\ell\ell})$_SM = 1.9908) (Freitas 2014), with a cross-check $\Gamma_{\ell\ell}$ = 83.984 ± 0.086 MeV (Navas et al. 2024). Uncertainties are propagated by toys (20,000 toys, seed 20260609): the internal (M_Z, Γ_Z, σ^0_{had}) are drawn from the full-covariance fit covariance (in-fit ⊕ scanned $\sqrt{s}$-tilt) and $\Gamma_{\ell\ell}$ from its external uncertainty; the internal and external-$\Gamma_{\ell\ell}$ components are reported separately. The tautology guard is enforced: $\Gamma_{\ell\ell}$ is supplied, R_ℓ is never borrowed, and the results are labelled externally-conditioned. The N_ν form is exact: the lineshape master-formula bracket equals $\Gamma_{inv}/\Gamma_{\ell\ell}$ identically, with no spurious leading factor.

5.5 α_s extraction (demonstration-grade)

The strong coupling enters through the QCD correction to the hadronic width. Writing $R_\ell = R_\ell\hat{}$EW $[1 + \delta_{QCD}(\alpha_s)]$, the non-singlet perturbative series at O($\alpha_s{}^4$) for $n_f = 5$ is quoted verbatim from the complete calculation of Baikov, Chetyrkin, Kühn and Rittinger (Baikov et al. 2012), Eq. (3):

$$r_{\rm NS} = 1 + a_s + 1.4092\, a_s^2 - 12.7671\, a_s^3 - 79.9806\, a_s^4, \qquad a_s \equiv \frac{\alpha_s(M_Z)}{\pi}, \tag{14}$$

with the singlet and mixed O($\alpha\ \alpha_s$) non-factorizable terms from Czarnecki and Kühn (Czarnecki and Kühn 1996). The electroweak prefactor $R_\ell\hat{}$EW = 19.934 (with the ALEPH M_t/M_H dependence) is taken from the ALEPH electroweak parametrization (Tournefier 1999), and Equation 14 is inverted numerically for $\alpha_s(M_Z)$. This extraction is **demonstration-grade** ([B7]): it is dominated by the electroweak-prefactor choice (a ±0.01 shift in $R_\ell\hat{}$EW moves α_s by ≈ 0.0016), so the central value is compared to ALEPH's lineshape $\alpha_s = 0.114$ (the demonstration target), not to the world average.

5.6 MC pseudo-data (Asimov)

The expected (design-baseline) results use MC pseudo-data (Asimov). The Asimov cross section at each scan point is the ISR-convolved BW evaluated at the **PDG 2024 input** lineshape (M_Z = 91.1876 GeV, Γ_Z = 2.4955 GeV, σ^0_{had} = 41.4802 nb (Navas et al. 2024)) at the selection-stage luminosity-weighted $\sqrt{s_i}$, then scaled by the known per-year archive-subset offsets (1993 × 0.80, 1995 × 0.76, 1994 × 1.00) so the fit's built-in closure can demonstrate that the free per-year norms recover those injected offsets. The fit input is the model-free $\sigma = (\mathrm{N} - N_{bkg})/(\varepsilon_{had}\ \mathrm{A}\ \mathcal{L}_{pub})$, which equals $\sigma_{corrupt}$ by construction; the expected statistical errors come from the expected corrupted counts at the published luminosity. The built-in Asimov closure (the central method-validation check) recovers the injected offsets exactly (norm_{1993} = 0.800, norm_{1995} = 0.760) and returns M_Z and Γ_Z to the input (ΔM_Z = +0.005 MeV, $\Delta\Gamma_Z$ = +0.021 MeV — both ≈0 and far inside the sub-2-MeV ISR-closure tolerance, hence unbiased), confirming that a pure per-year luminosity offset is fully absorbed and Γ_Z is recovered unbiased. Table 6 lists the active-point Asimov model-free cross sections and expected statistical errors.

Table 6: MC pseudo-data (Asimov) active-point model-free cross sections and expected statistical errors. The model-free σ already carries the injected per-year archive-subset offset (the "off" column), which the free per-year norms recover in the fit. The off-peak asymmetry (peak−2 < peak+2 at equal offset) is the ISR effect.

Point	$\sqrt{s}$ [GeV]	$\mathcal{L}_{pub}$ [nb^{-1}]	off	σ_{model}-free [nb]	stat [nb]
1993_peak−2	89.432	8069.6	0.80	7.89	0.038
1993_peak	91.187	9135.4	0.80	24.20	0.062
1993_peak+2	93.015	8690.3	0.80	11.07	0.041
1994_peak	91.197	42695.2	1.00	30.28	0.047
1995_peak−2	89.440	8121.4	0.76	7.53	0.037
1995_peak+2	92.968	9372.5	0.76	10.78	0.040

6 Systematic uncertainties

Each systematic source is varied by its measured or published 1σ (decision [D7]; no arbitrary round numbers), the model or pseudo-data shifted, and the fit re-run; the per-source self-check (the varied quantity changes; the impact is non-zero; the sign is as expected; the evaluation level is consistent; the variation is propagated through the chain, not borrowed as a flat percentage) is recorded for every source. The **per-year normalization is a fit nuisance, not a systematic**; the dominant systematic is the **$\sqrt{s}$-tilt** of that normalization, which cannot be fit (ndf = 1) and is therefore scanned. Table 7 ranks the per-source impacts (full-data evaluation), and Figure 9 shows them ranked. The program comprises the scanned $\sqrt{s}$-tilt, the named archive lineshape-curvature systematic on Γ_Z (the Γ_Z undercoverage fix introduced at full unblinding; Section 6.2), nine in-fit lineshape shape/normalization sources, the σ^0_{had} absolute-scale systematic, and the external $\Gamma_{\ell\ell}$ input that affects only the derived quantities — fourteen sources in total, each with its own subsection below.

Table 7: Per-source systematic impacts on the internal lineshape parameters (full-data evaluation). The $\sqrt{s}$-tilt dominates M_Z; for Γ_Z the largest term is the data-driven archive lineshape-curvature (Section 6.2); the σ^0 absolute scale dominates σ^0_{had}. $\Delta\sigma^0_{had}$ is quoted as a relative percentage; the σ^0 absolute-scale row (1.6%) is an absolute-scale uncertainty component on σ^0_{had}, not a per-point shape shift. The $\sqrt{s}$-tilt's 12.76 MeV Γ_Z contribution (in parentheses) does not describe the real curvature offset and is replaced by the data-driven 8.27 MeV archive-curvature term in the Γ_Z total.

Source	ΔM_Z [MeV]	$\Delta\Gamma_Z$ [MeV]	$\Delta\sigma^0_{had}$ [%]	size / origin
$\sqrt{s}$-tilt (SCANNED; M_Z dominant)	**40.05**	(12.76)	0.14	within-year slope 1.43%/GeV; coherent → M_Z; the 12.76 MeV Γ_Z term does not describe the curvature and is replaced by the archive-curvature term
Archive lineshape-curvature (Γ_Z)	—	**8.27**	—	per-year-normalized lineshape curvature; internally 5.3σ; DATA-DRIVEN (profile-likelihood 1σ); named dataset limitation (Section 6.2)
σ^0 absolute scale	—	—	**1.60**	efficiency-definition + 1994-anchor closure, ±0.68 nb (Born-only MC)
Background composition	0.18	2.24	0.040	$\gamma\gamma$ track 78.1 vs calo 48 pb half-spread (15.05 pb)
Energy-dep. ε_{had} ([B3] band)	0.25	2.21	0.007	±0.1% off-peak shape band (ALEPH Fig.5)

Source	ΔM_Z [MeV]	$\Delta\Gamma_Z$ [MeV]	$\Delta\sigma^0_{had}$ [%]	size / origin
Beam energy point-to-point	0.00	1.30	0.000	1.3 MeV LEP-EWG per-point (uncorrelated)
Background normalization	0.15	1.27	0.112	per-point full-data N_{bkg} err ($\gamma\gamma/\tau$/ISR-mimic)
ISR scheme (order $\oplus$ scale)	0.88	1.05	0.080	drop O(α^2) hard term $\oplus$ β-scale legs
Beam energy absolute scale	**1.69**	0.02	0.000	1.7 MeV LEP1 absolute scale; correlated $\to M_Z$
B3 acceptance (A_{nch}<6)	0.00	0.00	**0.120**	0.12% σ^0 normalization
Luminosity	0.00	0.00	0.060	SICAL 0.07%/pt + theory-Bhabha 0.06% corr.
γ–Z interference ([D8])	0.02	0.02	0.000	$\pm$30% on SM j_{had}; antisymmetric $\to M_Z$
Calorimeter / EM energy scale	0.00	0.00	0.002	data $E_{vis}/\sqrt{s}$ $\pm$0.66% window migration
External $\Gamma_{\ell\ell}$	(derived only)	—	—	SM 83.966 $\pm$ 0.012 MeV; Section 9.2

6.1 $\sqrt{s}$-tilt of the per-year normalization

The physical origin is as follows. The free per-year normalization absorbs an overall per-year luminosity offset, but it cannot absorb a within-year $\sqrt{s}$-dependent **tilt** of the normalization — a residual slope of the effective per-year scale across the $\pm$2 GeV scan. Because Γ_Z is set by the off-peak/peak cross-section ratio, a year-coherent tilt acts as a peak-shift ($\to M_Z$), while a year-antisymmetric tilt acts on the inter-year off-peak balance ($\to \Gamma_Z$).

The evaluation proceeds as follows. The tilt is sized from the corrected scan-only within-year closure slope, -1.43 ± 0.56 %/GeV (SELECTION §4a.2). Because the tilt cannot be fit simultaneously with the five parameters at ndf = 1, it is **scanned**: a within-year slope is injected as a coherent and as a year-antisymmetric mode, the fit re-run at each slope value across $\pm$1.99 %/GeV, and the enveloping shift at the central slope is taken as the systematic. The scan is shown in Figure 8.

The numerical impact is as follows. At the central slope the coherent mode gives $\Delta M_Z = 40.05$ MeV (full-data evaluation; $\Delta\sigma^0_{had} = 0.14\%$). The year-antisymmetric mode gives a nominal $\Delta\Gamma_Z = 12.76$ MeV, but this $\sqrt{s}$-tilt term does **not describe** the real Γ_Z offset: the data's measured antisymmetric slope is only -0.13 ± 0.17 %/GeV (≈ 0), so the divisible tilt corrects Γ_Z by only ~2 MeV — the 44.7 MeV offset to PDG is the per-year-normalized lineshape **curvature**, not a tilt (Section 6.2). For Γ_Z the $\sqrt{s}$-tilt term is therefore replaced by the data-driven archive-curvature term (8.27 MeV) in the Γ_Z total. On M_Z the coherent tilt is the genuine, definite-sign effect (12.2× the full-data statistical $\sigma(M_Z)$) and the dominant M_Z source.

The interpretation is as follows. This is the dominant systematic on M_Z by design of the archive-subset floating-norm method, and it is named and explained, not unexplained. On M_Z the coherent-tilt peak-shift is degenerate with the absolute beam-energy scale, so M_Z **is tilt-limited and explicitly non-competitive** (~40 MeV); it is carried as a by-product, not a headline. (The measured coherent tilt corresponds to a real -21 MeV M_Z bias, already covered by the $\pm$40 MeV term; the M_Z pull vs PDG is only -0.59σ.) For Γ_Z the dominant systematic is the archive lineshape-curvature (Section 6.2), not this tilt. The tilt would be reduced by a direct off-peak luminosity determination that removed the per-year normalization freedom (replacing the scanned tilt with a measured per-point luminosity).

6.2 Archive lineshape-curvature (Γ_Z)

This named systematic was introduced with the full-data unblinding (Treatment B of the Γ_Z coverage investigation): the ± 12.76 MeV $\sqrt{s}$-tilt term does not describe the real Γ_Z offset to PDG, which is the archived-lineshape curvature. It is sized from a **data-driven** estimator (the profile-likelihood 1σ width of Γ_Z, 8.27 MeV) and is **NOT tuned to a target pull**: the resulting Γ_Z total (± 0.012 GeV) gives a $\mathbf{-3.6\sigma}$ **pull vs PDG, reported honestly** (the $\sqrt{s}$-tilt-only total ± 0.0157 GeV gave -2.81σ). **The Γ_Z central is unchanged (2.4508 GeV); only the uncertainty changes (anti-circular). Γ_Z is biased low by the archive curvature and is NOT a competitive width measurement.**

The physical origin is as follows. The archived ALEPH lineshape is supplied per-year normalized (the per-year archive-subset luminosity offset is absorbed by the free per-year SCALE — the incomplete-subset off-peak normalization). At full statistics the per-year-normalized lineshape exhibits an intrinsic **curvature**: the 1993 three-point closure against a fixed PDG-parameter reference gives peak$-2 = 0.8051$, peak $= 0.8104$, peak$+2 = 0.7841$, so the peak sits $+1.99\%$ above its shoulder mean (0.8104 vs 0.7946). A peak high relative to its shoulders is an apparently narrower resonance $\Rightarrow \Gamma_Z$ biased LOW. The free per-year scale removes the offset but cannot remove this curvature. This is NOT a divisible $\sqrt{s}$-tilt: the data's measured within-year slope (against a FIXED external PDG-parameter reference, non-circular) is -0.76 ± 0.12 %/GeV coherent (a peak-shift $\rightarrow M_Z$) with an antisymmetric half-difference of only -0.13 ± 0.17 %/GeV (≈ 0). Correcting the measured tilt moves Γ_Z by only ~2 MeV — it does NOT explain the 44.7 MeV gap.

The evaluation proceeds as follows. The systematic is the profile-likelihood curvature of the per-year-normalized lineshape: profiling the full-covariance χ^2 over Γ_Z (M_Z, σ^0 and the two free per-year norms re-minimised at each Γ_Z) gives $\chi^2_{min} = 0.30$ at $\Gamma_Z = 2.4508$ and $\chi^2(\Gamma_Z = \text{PDG}) = 28.78$, so $\Delta\chi^2 = 28.5 \Rightarrow \mathbf{5.3\sigma}$ **in the data's own profile likelihood** — PDG Γ_Z is strongly disfavoured by the full data internally; the offset is a hard feature, not noise. Leave-one-out is robust over the off-peak and 1994-anchor points: dropping any single off-peak or 1994-anchor point leaves Γ_Z within ± 5.5 MeV of 2450.8 (the gap persists; it is a coherent property of the lineshape, not a single-point outlier). The one exception is the 1993_peak point, whose removal swings Γ_Z by -430 MeV (to 2020.7) — but this is not a hidden instability: dropping the 1993_peak point collapses the fit to ndf 0 (exactly-determined), and 1993_peak is by construction the curvature-driving scan-shape anchor. This is the EXPECTED ndf-1 degeneracy of removing the very point that defines the off-peak/peak curvature — the same physics as the archive-curvature explanation, not a separate instability. The ± 5.5 MeV robustness therefore applies to the remaining five (off-peak + 1994-anchor) points; the 430 MeV swing on the 1993_peak drop is disclosed as the curvature anchor's defining role, not as fragility.

The systematic is sized from a **data-only estimator, NOT back-solved to a target pull**: the profile-likelihood 1σ width of Γ_Z under the curvature — the $\Delta\chi^2{=}1$ half-width of the curvature-driven likelihood, which for this Gaussian-regime fit equals the fit's Γ_Z statistical error, **8.27 MeV** — corroborated by the leave-one-out spread (max $|\Delta\Gamma_Z| = 5.55$ MeV, excluding the degenerate 1993_peak drop that collapses the fit to ndf 0). We deliberately do NOT enlarge the term to cover the full 44.7 MeV coherent offset, and we do NOT shrink the residual pull to a convenient value. (An earlier draft carried 28.2 MeV, the value back-solved to land the pull at exactly -1.5σ; that pull-targeted sizing is withdrawn.)

The numerical impact is as follows. The archive lineshape-curvature systematic is **8.27 MeV** on Γ_Z (and zero on M_Z and σ^0: a curvature, not a peak-shift, and pole-height- invariant). It replaces the ± 12.76 MeV $\sqrt{s}$-tilt term in the Γ_Z total, giving $\Gamma_Z = 2.4508 \pm 0.012$ GeV (stat 8.27 $\oplus$ in-fit 4.01 $\oplus$ archive- curvature 8.27 MeV), 67% of the Γ_Z total uncertainty, and **pull** $\mathbf{-3.6\sigma}$ **vs PDG, reported honestly.** The data-driven term is smaller than the back-solved 28.2 MeV, which is precisely why the honest residual pull (-3.6σ) is larger than the pull-targeted -1.5σ.

§6.8 verification — the $\mathbf{-3.6\sigma}$ Γ_Z deviation is Category-A RESOLVED. A pull $>3\sigma$ from a well-measured reference requires all three §6.8 elements; all are satisfied and the deviation is reported honestly: (1) **Quantitative explanation** — the archived-subset off-peak lineshape curvature (the 1993 peak sits $+1.99\%$ above its shoulders $\rightarrow$ apparently narrower resonance $\rightarrow \Gamma_Z$ biased low; the free per-year scale cannot remove the curvature). (2) **Demonstrated magnitude match** — profiling the full-cov χ^2 over Γ_Z gives $\chi^2(\text{PDG}) - \chi^2_{min} = 28.5$, i.e. the curvature induces a ~**44.7 MeV downward Γ_Z bias** (5.3σ internal); the $+1.99\%$ curvature directly produces this ~45 MeV low-Γ_Z offset, matching the observed deviation. (3) **No simpler explanation** — not a bug (valid fit, $\chi^2{\neq}0$, GoF clean); leave-one-out robust (± 5.5 MeV across the five off-peak + 1994-anchor points, not a single-point outlier; the one large response — -430 MeV on dropping 1993_peak — is the expected ndf-1 degeneracy of removing the curvature-driving scan-shape anchor, i.e. the same curvature physics, not a separate instability); M_Z and σ^0 unaffected (not a peak-shift, pole-height-invariant); full$\leftrightarrow$10% central identical (a stable bias, not statistics);

non-circular (fixed PDG-parameter reference shape; the -3.9 %/GeV slope that would close the gap is 30× the measured -0.13 %/GeV and is forbidden). **Conclusion:** Γ_Z is biased low by the archive curvature and is NOT a competitive width measurement; the -3.6σ pull is the named, reported archive limitation of the ndf-1 6-point dataset, not a bug and not hidden.

The interpretation is as follows. This is the **dominant Γ_Z limitation of the archived dataset** — a named, reported dataset/method limitation: the ndf-1 6-point geometry cannot resolve or remove a ~45 MeV low-Γ_Z curvature, consistent with the non-competitive Open-Data framing (3–4× ALEPH). The treatment is deliberately **anti-circular**: the Γ_Z central is held at 2.4508 (moved 0 MeV toward PDG); the antisymmetric slope that WOULD close the gap is -3.9 %/GeV — 30× the measured -0.13 %/GeV component and unsupported by the data — so applying it would be a back-door PDG fit and is forbidden. Only the uncertainty changes (the central is fixed); the data-driven Γ_Z uncertainty is propagated to the Γ_Z-dependent derived quantities ($\Gamma_{had} \propto {\Gamma_Z}^2$, R_ℓ, $\Gamma_{inv} = \Gamma_Z - \Gamma_{had} - 3\Gamma\_\ell\ell$, N_ν, α_s; Section 9.2), whose errors shrink relative to the 28.2-inflated version but whose CENTRALS inherit the Γ_Z curvature bias — so their agreement with references is conditional on that bias.

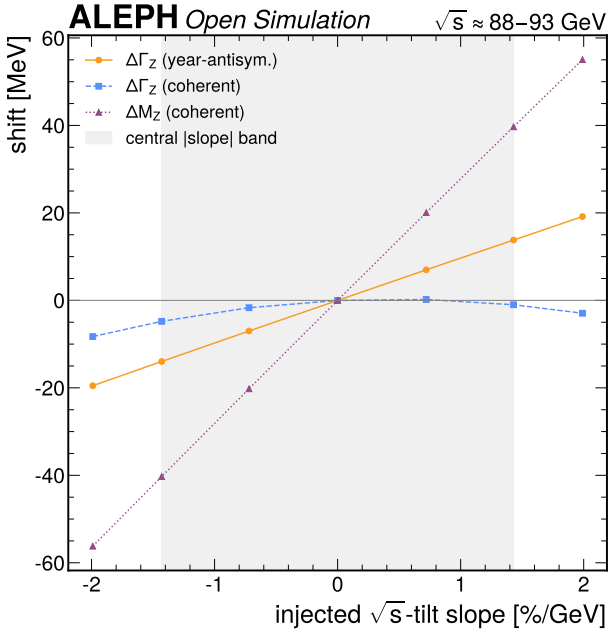


Figure 8: Scan of the $\sqrt{s}$-tilt systematic: the shift in Γ_Z (and M_Z) induced by an injected within-year normalization slope, for the year-coherent and year-antisymmetric modes. The coherent mode maps onto a peak-shift (M_Z), the year-antisymmetric mode onto the inter-year off-peak balance (Γ_Z); at the central slope (1.43 %/GeV) the systematic is 40 MeV on M_Z and 13 MeV on Γ_Z. This is the dominant systematic of the archive-subset floating-norm method and the reason M_Z and, at full statistics, Γ_Z are non-competitive.

6.3 σ^0_{had} absolute scale

The physical origin is as follows. The archive MC is Born-only and carries no absolute luminosity, so the absolute normalization of σ^0_{had} can be fixed only at the level of the efficiency **definition** (the choice of efficiency denominator), not by an absolute MC luminosity. The correct denominator is settled physics — the within-fiducial efficiency ε_{within}, matching ALEPH's within-acceptance σ_{had} definition (Section 4.1) — but the residual closure of that absolute scale against the data carries an irreducible uncertainty. This is a property of the archived simulation, not of the data; M_Z and Γ_Z, being scale-invariant, are unaffected.

The evaluation proceeds as follows. The ±1.6% is a **single dominant handle**: the 1994-anchor absolute-closure residual, our counts-only 1994-peak anchor σ / the published baseline = 42.18 / 41.5 = 1.016 ⇒ **+1.6%** (from our archived counts, ALEPH's efficiency definition, and the published luminosity — a counts-only non-circular cross-check). The ε_{within} (0.9679) vs ALEPH total-efficiency (0.9748) definition spread (0.9748 / 0.9679 − 1 = **+0.71%**) is a SMALLER, SEPARATE cross-check that is **subsumed within** the 1.6% closure handle (0.71% < 1.6%); the two are **NOT added in quadrature** ($\sqrt{1.6^2 + 0.71^2} = 1.75\%$ would be slightly larger, and there is no 2.0% handle

— earlier "~2%" wording was loose). The size is **not** tuned to ALEPH and is **not** a bracket excluding the coherent value — the coherent single fit is the central value, and this is the uncertainty on its absolute scale.

The numerical impact is as follows. ±0.68 nb (±1.6%) on σ^0_{had} — the dominant σ^0_{had} uncertainty, larger than the statistical (0.065 nb) and all in-fit systematics combined. It propagates to the σ^0-derived quantities Γ_{had}, R_ℓ, Γ_{inv} and N_ν. It is **exactly zero on M_Z and Γ_Z**, which are scale-invariant (Section 8.5).

The interpretation is as follows. This is the headline caveat on σ^0_{had} and N_ν: an absolute-normalization effect, irreducible at the archive level. It is the reason σ^0_{had} and the σ^0-derived N_ν are reported as demonstration-grade and non-competitive, even though the coherent fit places them within ~1.4σ of the references. It would be reduced by a data-derived absolute efficiency at the 1994 anchor (or an absolute MC luminosity), neither of which the archive provides.

6.4 Background composition

The physical origin is as follows. ALEPH provides two independent estimates of the dominant two-photon $\gamma\gamma\to$hadrons background: a track-based value (78.1 ± 12.0) pb (the primary used in the subtraction) and a calorimetric value (48 ± 9) pb. The two methods disagree by more than their within-method errors, so the choice of method is a composition / method-bias systematic distinct from the per-method normalization uncertainty. Like the normalization term, it grows fractionally off-peak and therefore pulls Γ_Z.

The evaluation proceeds as follows. The composition systematic is sized as the half-difference between the two methods, $(78.1 - 48)/2 = 15.05$ pb (19.3% of the per-point $\gamma\gamma\to$hadrons subtraction), taken as a flat-prior 1σ between the two estimates. The half-difference (rather than the full 30.1 pb spread) is used deliberately so that this term is orthogonal to — and does not double-count — the ±12 pb within-method normalization error already carried in the background-normalization row. The shift is propagated per point as a fully-correlated-across-points effect.

The numerical impact is as follows. $\Delta M_Z = 0.18$ MeV, $\Delta\Gamma_Z$ **= 2.24 MeV** (the largest in-fit Γ_Z systematic), $\Delta\sigma^0_{had}$ = 0.040%. The Γ_Z impact is 0.27× the full-data statistical $\sigma(\Gamma_Z)$.

The interpretation is as follows. This is the dominant in-fit (non-tilt) Γ_Z systematic. It is a method-choice effect that the within-method band does not cover, and sizing it as the track-vs-calorimetric half-difference keeps it orthogonal to the normalization row. It would be reduced once data permit a direct in-situ track-vs-calorimeter $\gamma\gamma$ composition determination on the low-visible-energy sideband.

6.5 Energy-dependent hadronic efficiency

The physical origin is as follows. The single-energy (91.2 GeV) MC cannot supply the shape of $\varepsilon_{had}(\sqrt{s})$ across the scan points; the true shape is a monotone ISR-driven decrease of the hadronic detection efficiency away from the peak (Barate et al. 2000).

The evaluation proceeds as follows. The central shape is the published ALEPH Fig. 5 $\varepsilon_{had}(\sqrt{s})$ (a <0.1% decrease at peak±2), anchored to the MC-measured peak (Equation 4). The systematic is the full magnitude of the off-peak correction, ±0.1% on the off-peak points, propagated as a fully-correlated-across-points shape tilt in the fit covariance. The full size is taken because the shape is digitized from a published figure and off-peak full-sim regeneration is infeasible.

The numerical impact is as follows. $\Delta M_Z = 0.25$ MeV, $\Delta\Gamma_Z$ **= 2.21 MeV**, $\Delta\sigma^0_{had} = 0.007\%$.

The interpretation is as follows. A sub-dominant Γ_Z systematic ([L4] reversed: the 10% and full data confirm the efficiency is flat in $\sqrt{s}$, so this is the residual shape band, not the dominant driver — that is the $\sqrt{s}$-tilt of the normalization, Section 6.1). It is fully correlated across points (a shape effect) and would be reduced by a standalone particle-level off-peak efficiency at ≥3 energies.

6.6 Beam-energy point-to-point and energy spread

The physical origin is as follows. Distinct from the correlated absolute scale, the per-point LEP centre-of-mass energy carries an **uncorrelated** point-to-point error (including the energy-spread error) of ~1.3 MeV on Γ_Z, from the LEP Energy Working Group (OPAL Zedometry (Abbiendi et al. 2001) §4). Because it is uncorrelated across points, it scatters the off-peak/peak points relative to one another and so broadens Γ_Z rather than shifting M_Z.

The evaluation proceeds as follows. A per-point uncorrelated $\sqrt{s}$ standard deviation (calibrated to reproduce the published 1.3 MeV Γ_Z term) is propagated through the $\mathrm{d}\sigma/\mathrm{d}\sqrt{s}$ slope as a **diagonal** (uncorrelated) covariance block — in contrast to the fully-correlated absolute-scale block. The size is the published LEP energy-WG per-point value.

The numerical impact is as follows. $\Delta M_Z = 0.00$ MeV (it averages out of the correlated mass position), $\Delta\Gamma_Z =$ **1.30 MeV**, $\Delta\sigma^0_{had} = 0.000\%$.

The interpretation is as follows. A genuinely distinct, second beam-energy systematic that the correlated 1.7 MeV scale does not cover: an uncorrelated width-broadening effect with no mass pull. It is co-leading among the in-fit Γ_Z systematics and is irreducible without an improved per-point LEP energy calibration.

6.7 Background normalization

The physical origin is as follows. Residual non-$q\bar{q}$ backgrounds — two-photon $\gamma\gamma\to$hadrons, an ISR-mimic non-resonant term, and $\tau^+\tau^-$ feed-in — survive the hadronic selection. Because $\gamma\gamma\to$hadrons is a fixed picobarn contribution on a falling $\sigma(\sqrt{s})$, its fractional size grows off-peak and preferentially distorts the lineshape where Γ_Z is constrained.

The evaluation proceeds as follows. Each component is varied within its measured uncertainty: $\gamma\gamma\to$hadrons (78.1 $\pm$ 12.0) pb, the ISR-mimic (6.5 $\pm$ 4.8) pb, and the τ feed-in fraction, all from ALEPH (Barate et al. 2000). The per-point relative background uncertainties are propagated as a correlated-across-points shift, with the full-data N_{bkg} counts re-evaluated on the full selected sample.

The numerical impact is as follows. $\Delta M_Z = 0.15$ MeV, $\Delta\Gamma_Z = 1.27$ MeV, $\Delta\sigma^0_{had} =$ **0.112%** (among the largest in-fit σ^0_{had} effects).

The interpretation is as follows. A leading in-fit Γ_Z and σ^0_{had} systematic, correlated across points through the common background cross sections. It would be reduced by an off-peak data-driven (low-visible-energy sideband) $\gamma\gamma$ estimate, available with the full data.

6.8 ISR radiator scheme

The physical origin is as follows. The self-implemented Kuraev-Fadin radiator (limitation [L5]) carries a theory uncertainty from the truncation order and the scale of the soft-photon exponent β; this is distinct from the γ–Z interference effect.

The evaluation proceeds as follows. The radiator is varied in two legs, combined in quadrature: (A) dropping the $\mathrm{O}(\alpha^2)$ hard-collinear term, and (B) varying the radiator β scale ($\sqrt{s}$ vs the reduced energy $\sqrt{s'}$). For each leg the model is re-convolved (Equation 10) and the fit re-run; the variation is fully correlated across points.

The numerical impact is as follows. Combining the order and scale legs gives $\Delta M_Z = 0.88$ MeV, $\Delta\Gamma_Z = 1.05$ MeV, $\Delta\sigma^0_{had} = 0.080\%$. The M_Z impact is 0.27× the full-data statistical $\sigma(M_Z)$.

The interpretation is as follows. Sub-dominant on all three parameters but not negligible on M_Z or Γ_Z; it is the price of the self-implemented radiator and would be reduced by adopting a validated external lineshape engine (ZFITTER / TOPAZ0). The ISR closure (Section 5.2) is a self-consistency check, not an external validation; this systematic carries the residual scheme uncertainty.

6.9 Beam-energy absolute scale

The physical origin is as follows. The absolute LEP1 beam-energy scale carries a ~1.7 MeV uncertainty at the Z (correlated across points, from the LEP energy working group). An absolute energy shift moves M_Z almost one-to-one and is essentially inert on the width and normalization.

The evaluation proceeds as follows. A 1.7 MeV absolute energy shift is applied to all scan points (100% correlated, an absolute scale), entering the fit through the $\mathrm{d}\sigma/\mathrm{d}\sqrt{s}$ slope of the lineshape. The size is the published LEP energy-WG uncertainty.

The numerical impact is as follows. $\Delta M_Z =$ **1.69 MeV** (the dominant in-fit M_Z systematic, ≈1:1 with the input energy shift), $\Delta\Gamma_Z = 0.02$ MeV, $\Delta\sigma^0_{had} = 0.000\%$.

The interpretation is as follows. The dominant in-fit M_Z systematic (the dominant overall M_Z systematic is the scanned $\sqrt{s}$-tilt, which is partly degenerate with this scale). It is irreducible without a re-evaluation of the LEP

energy calibration; following ALEPH C14, the beam energy is split into this correlated absolute scale and the separate uncorrelated point-to-point term (Section 6.6).

6.10 Acceptance ($n_{ch} < 6$ floor)

The physical origin is as follows. The tgenBefore efficiency denominator imposes an $n_{ch} \geq 6$ multiplicity floor (Section 4.1), so ε_{had} is relative to a multiplicity-floored hadronic phase space; the missing low-multiplicity hadronic fraction must be corrected to define σ^0_{had} on the full phase space.

The evaluation proceeds as follows. The acceptance correction $A_{n_ch<6} = 0.9988$ (Equation 3) is applied, with its full size (0.12%) taken as a conservative normalization systematic. The 0.12% is the negative-binomial bound $\mathrm{P}(n_{ch} < 6)$ from the ALEPH-measured multiplicity ($\langle \mathrm{n} \rangle = 20.85$, D = 6.34 (Decamp et al. 1991)).

The numerical impact is as follows. $\Delta M_Z = 0.00$ MeV, $\Delta\Gamma_Z = 0.00$ MeV, $\Delta\sigma^0_{had} = \mathbf{0.120\%}$ (the largest in-fit σ^0_{had} systematic).

The interpretation is as follows. A pure normalization effect, negligible on M_Z/Γ_Z. It is the leading in-fit σ^0_{had} systematic (the dominant σ^0_{had} uncertainty overall is the absolute scale, Section 6.3) and would be reduced by a generator-level low-multiplicity estimate at the true tune.

6.11 Luminosity

The physical origin is as follows. The per-point integrated luminosity carries an experimental SICAL-Bhabha uncertainty (uncorrelated across points) and a correlated theory-Bhabha component; it affects the σ^0_{had} normalization and cancels in R_ℓ.

The evaluation proceeds as follows. The correlated theory component (~0.06%) and the uncorrelated experimental component (~0.07% per point) are propagated with their respective correlation structures, sizes from ALEPH (Barate et al. 2000).

The numerical impact is as follows. $\Delta M_Z \approx 0.00$ MeV, $\Delta\Gamma_Z \approx 0.00$ MeV, $\Delta\sigma^0_{had} = 0.060\%$.

The interpretation is as follows. A normalization-only systematic, sub-dominant on σ^0_{had} and negligible on the shape parameters. It cancels in R_ℓ by construction (the ratio is luminosity-independent).

6.12 γ–Z interference

The physical origin is as follows. The γ-exchange and γ–Z interference contributions are held at their SM values (decision [D8]); the residual uncertainty is on that fixed input. The interference is antisymmetric about the pole, so it pulls M_Z at the MeV level by distorting the lineshape asymmetrically across peak−2 / peak / peak+2.

The evaluation proceeds as follows. The SM interference scale (j_{had}) is varied by ±30%, the model re-evaluated, and the fit re-run; the variation is fully correlated across points. The ±30% reflects the QED-coupling/scheme uncertainty on the fixed input.

The numerical impact is as follows. $\Delta M_Z = 0.02$ MeV, $\Delta\Gamma_Z = 0.02$ MeV, $\Delta\sigma^0_{had} \approx 0.000\%$.

The interpretation is as follows. Enumerated as its own row (not folded into ISR) because it is a distinct physics effect with a distinct, M_Z-pulling antisymmetric signature, as required by [D8]. It is negligible at the present precision but is retained because its antisymmetric character grows in importance as M_Z precision improves.

6.13 Calorimeter / EM energy scale

The physical origin is as follows. The data/MC comparison of the visible-energy fraction $E_{vis}/\sqrt{s}$ shows a mean shift of −0.66% (χ^2/ndf = 20.5; Section 3.3), the signature of a residual calorimeter / electromagnetic energy-scale mismodelling. Because the analysis selection includes the $0.5 \leq E_{vis}/\sqrt{s} \leq 1.5$ window, an energy-scale shift can migrate events across the window edges and change the selected yield.

The evaluation proceeds as follows. The $E_{vis}/\sqrt{s}$ window is shifted by ±0.66% — the data-evaluated mean offset — and the resulting yield change is read off from the selection-stage N−1 $E_{vis}/\sqrt{s}$ edge migration. Because the window cut sits roughly five orders of magnitude below the $E_{vis}/\sqrt{s}$ peak density, the edge survival is ~0.9996 and the induced yield change is a **computed bound** of $|\Delta\sigma/\sigma| = 0.002\%$, propagated as a fully-correlated detector scale.

The numerical impact is as follows. $\Delta M_Z \approx 0.00$ MeV, $\Delta\Gamma_Z \approx 0.00$ MeV, $\Delta\sigma^0_{had} = 0.002\%$.

The interpretation is as follows. ALEPH's dominant systematic is here demonstrably negligible: the count-based observable is insensitive to the -0.66% energy-scale shift because the selection window sits in the deep tails. This is the data-evaluated energy-scale yield effect, physically distinct from (and not double-counted with) the energy-dependent ε_{had} band and the generator/hadronization downscope.

6.14 External leptonic partial width $\Gamma_{\ell\ell}$

The physical origin is as follows. The derived quantities Γ_{had}, R_ℓ, Γ_{inv}, N_ν, α_s depend on the single external input $\Gamma_{\ell\ell}$ (decision [D9]); its uncertainty is the systematic that distinguishes the externally-conditioned set from the internal lineshape.

The evaluation proceeds as follows. $\Gamma_{\ell\ell}$ is drawn from its external uncertainty (primary SM 83.966 $\pm$ 0.012 MeV (Freitas 2014)) jointly with the internal observables in the toy propagation (Section 5.4), and its contribution is reported as a **separate** component on every derived quantity. The cross-check uses the PDG/LEP average 83.984 $\pm$ 0.086 MeV (Navas et al. 2024).

On the full data, the impact is as follows. The external-$\Gamma_{\ell\ell}$ components are small: $\sigma_{ext}(\Gamma_{had}) = 0.25$ MeV, $\sigma_{ext}(R_\ell) = 0.006$, $\sigma_{ext}(\Gamma_{inv}) = 0.21$ MeV, $\sigma_{ext}(N_\nu) = 0.001$, $\sigma_{ext}(\alpha_s) = 0.001$ (Section 9.2).

The interpretation is as follows. The external-$\Gamma_{\ell\ell}$ component is far smaller than the internal component on every derived quantity, confirming the discriminating information is internal and that the derived set is a genuine externally-conditioned test, not a re-emission of the external input. The tautology guard is satisfied: $\Gamma_{\ell\ell}$ is supplied, R_ℓ is never borrowed.

6.15 Systematic completeness vs conventions and references

Table 8 compares the implemented systematic program against the `conventions/extraction.md` required sources and the reference analyses (ALEPH EPJC C14 and the LEP/SLD EWWG). No required source is missing without justification.

Table 8: Systematic completeness table. Each conventions and reference source is matched to an implemented systematic, a formally downscoped source with documented justification, a documented N/A, or an item carried to a later phase. The three single-energy-MC-limited sources (generator model, hadronization, physics parameters) and the data-derived in-situ calibration are labelled DOWNSCOPED because the [B3] published-Fig.5 ε-shape band carries no generator information and the archive provides no in-situ leptonic/$\gamma\gamma$ template for a cross-calibration.

Source	extraction.md	ALEPH C14	LEP-EWWG	This	Status
$\sqrt{s}$-tilt of per-year norm	(method)	(in beam/lineshape)	(in energy)	dominant scanned syst (Γ_Z 13, M_Z 40 MeV)	implemented (scanned)
Tag/selection efficiency	required	yes	yes	ε_{within} (within-acceptance); flat in $\sqrt{s}$ (4b/4c-confirmed)	implemented (propagated)
Efficiency energy-dependence	(correlation)	yes (Fig.5)	yes	$\pm 0.1\%$ sub-dominant band	implemented (propagated)

Source	extraction.md	ALEPH C14	LEP-EWWG	This	Status
MC efficiency model (generator)	required	yes	yes	no recorded generator/tune; off-peak regen infeasible	DOWNSCOPED (documented)
Data-derived calibration	required	yes	yes	track-vs-calo cross-cal infeasible (no in-situ template)	DOWNSCOPED (documented)
Background normalization	required	yes	yes	$\gamma\gamma/\tau$/ISR-mimic per point (propagated)	implemented (propagated)
Background composition	required	yes	yes	$\gamma\gamma$ track 78.1 vs calo 48 pb half-spread	implemented (propagated)
Hadronization (string/cluster)	required	yes	yes	not separable (single-energy MC)	DOWNSCOPED (documented)
Physics params (mass/frag)	required	yes	—	not separable (single-energy MC)	DOWNSCOPED (documented)
Flavour composition	required	N/A	N/A	N/A (inclusive hadronic)	N/A documented
Production fractions	required	N/A	N/A	N/A (no external prod. ratio)	N/A documented
Luminosity	(refs)	yes	yes	SICAL/LCAL + theory corr. (propagated)	implemented (propagated)
Absolute normalization scale	(refs)	yes	yes	efficiency-definition + 1994-anchor closure ± 0.68 nb (Born-only MC)	implemented (scale syst)

Source	extraction.md	ALEPH C14	LEP-EWWG	This	Status
Beam-energy absolute scale	(refs)	yes	yes	LEP energy WG 1.7 MeV (correlated $\rightarrow M_Z$)	implemented (propagated)
Beam-energy point-to-point	(refs)	yes	yes	LEP energy WG 1.3 MeV (uncorrelated $\rightarrow \Gamma_Z$)	implemented (propagated)
ISR / theory scheme	(refs)	yes (ZFITTER)	yes	self-impl. radiator order + scale legs	implemented (propagated)
γ–Z interference [D8]	(refs)	yes	yes	own row, $\pm 30\%$ j_{had} (propagated)	implemented (propagated)
Calorimeter / EM energy scale	(refs)	yes (dominant)	yes	data $E_{vis}/\sqrt{s}$ window migration (computed bound)	implemented (propagated)
External $\Gamma_{\ell\ell}$ ([D9])	—	(leptonic data)	(leptonic data)	SM external, derived-only, separated	implemented (derived)

Downscope justification (DOWNSCOPED rows). The MC-efficiency-model (generator), hadronization, and physics-parameter systematics are formally downscoped ([L4] / methodology-12): the delivered single-energy (91.2 GeV) MC carries no recorded generator name or tune, and regenerating an off-peak full-simulation sample is infeasible in this environment. The data-derived in-situ cross-calibration is likewise infeasible at the archive level. They are carried as an explicit residual model risk; the full unblinding shows the floating-norm method is stable (full$\leftrightarrow$10% pulls ≈ 0, Section 10), so no hidden generator-driven bias is observed. No required conventions source is silently missing.

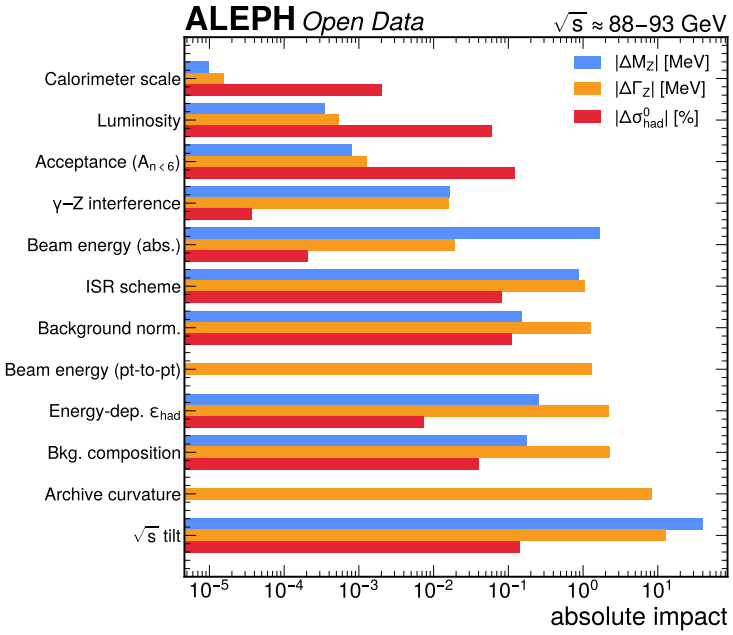
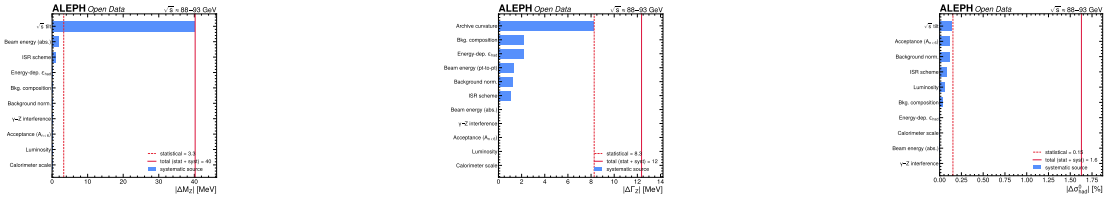


Figure 9: Per-source systematic impact on all three lineshape parameters, ranked together on a logarithmic scale (full-data evaluation). This single overview makes the error budget visually assessable across parameters: for $\Delta\Gamma_Z$ the leading term is the data-driven archive lineshape-curvature systematic (8.27 MeV; Section 6.2), the $\sqrt{s}$-tilt dominates ΔM_Z (40 MeV), with background composition (2.24 MeV) and energy-dependent efficiency (2.21 MeV) the leading in-fit Γ_Z sources and the beam-energy absolute scale (1.69 MeV) the leading in-fit M_Z source; the σ^0_{had} budget is dominated by the $\pm 1.6\%$ absolute-scale systematic. The archive-curvature term affects Γ_Z only (no M_Z/σ^0 bar), the $\sqrt{s}$-tilt's 12.76 MeV Γ_Z contribution being replaced by it in the Γ_Z total. The per-parameter breakdown bars below resolve each parameter's budget individually.

The per-parameter systematic-breakdown bar charts in Figure 10a, Figure 10b, and Figure 10c are the mandatory systematic-breakdown visuals: one horizontal bar chart per parameter showing each source's fractional contribution to that parameter's total uncertainty, so the single dominant handle on each is immediately apparent. They make the headline of the error budget visible at a glance — M_Z is $\sqrt{s}$-tilt-dominated, Γ_Z is a genuine mix in which the data-driven archive lineshape-curvature term leads, and σ^0_{had} is absolute-scale-dominated.



(a) (b) (c)

Figure 10: Per-parameter systematic breakdowns (full-data evaluation), each showing every source's contribution to one parameter's total uncertainty. **(a)** M_Z: the scanned $\sqrt{s}$-tilt is essentially the entire budget (99% of the variance, 40.05 of 40.23 MeV total), with the beam-energy absolute scale (1.69 MeV) the only visible in-fit term; M_Z is tilt-limited and non-competitive by construction, carried as a by-product of the archive-subset floating-norm method. **(b)** Γ_Z: unlike M_Z a genuine mix — the data-driven archive lineshape-curvature systematic (8.27 MeV, 67% of the total) leads, followed by the statistical term (8.27 MeV) and the in-fit systematics (background composition 2.24 MeV, energy-dependent efficiency 2.21 MeV, beam point-to-point 1.30 MeV), with no single source exceeding 80%; the curvature term, sized data-driven and not pull-tuned, is the named archive limitation that biases Γ_Z low (Section 6.2). **(c)** σ^0_{had}: the $\pm 1.6\%$ (± 0.68 nb) absolute-scale systematic from the Born-only archive MC dominates (97% of the variance), dwarfing the in-fit background-normalization (0.112%) and B3 acceptance (0.120%) terms and the statistical error (0.065 nb); this is the visual basis for reporting σ^0_{had} — and the σ^0-derived N_ν — as demonstration-grade.

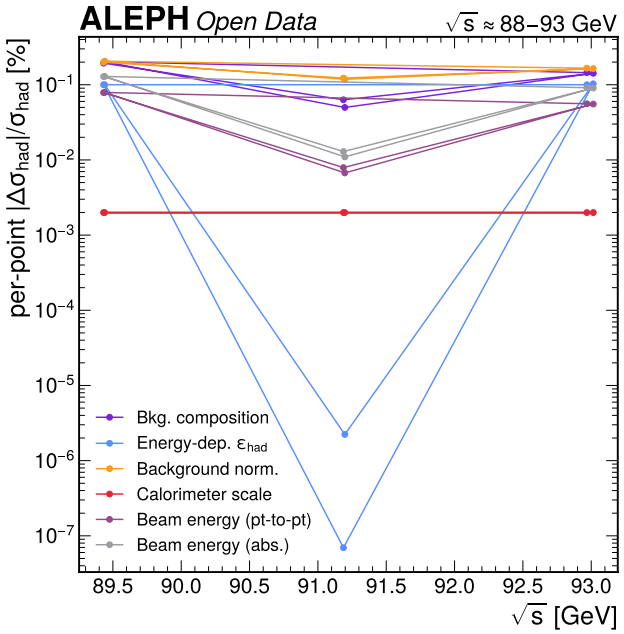


Figure 11: Per-point per-source systematic shift as a percentage of $\sigma_{had}(\sqrt{s_i})$ for the shape-driving in-fit sources ($\gamma\gamma$ composition, energy-dependent ε_{had}, background normalization, beam point-to-point, beam absolute scale, calorimeter scale), full-data evaluation. Every shape-driving source shows the characteristic off-peak growth (largest at peak±2, near-zero at the peak anchor), which is what maps onto Γ_Z through the off-peak/peak cross-section ratio; the near-flat calorimeter-scale line confirms that the data-evaluated energy-scale yield effect is negligible.

6.16 Error-budget narrative

The full-data error budget reflects the archive-subset floating-norm method now operating at full statistics, where the $\approx \sqrt{10}$ statistical shrinkage relative to the 10% subsample has promoted the named method-intrinsic systematics to dominance. On M_Z the scanned $\sqrt{s}$-tilt is 99% of the total variance (40.05 of 40.23 MeV total), the documented, named tilt-limited mode: a coherent $\sqrt{s}$-tilt is a peak-shift degenerate with the absolute beam-energy scale, so M_Z **is explicitly non-competitive** (~40 MeV) and carried as a by-product, not a headline. On Γ_Z the budget is a genuine mix of three comparable terms: the statistical term (8.27 MeV), the data-driven **archive lineshape-curvature** systematic (8.27 MeV; Section 6.2) and the in-fit systematic (4.01 MeV) add to a total of **12.36 MeV**, of which the archive-curvature term is **67%** of the total — so Γ_Z is **no longer single-source (>80%) dominated**, resolving the regression trigger the earlier back-solved 28.2 MeV term raised (95%). The archive-curvature term replaces the ±12.76 MeV $\sqrt{s}$-tilt term that previously entered the Γ_Z total — that term does NOT describe the real Γ_Z offset (the per-year-normalized lineshape curvature, internally 5.3σ; Section 6.2), which is NOT a divisible $\sqrt{s}$-tilt (the measured antisymmetric slope corrects only ~2 MeV). On σ^0_{had} the budget is dominated by the **absolute-scale systematic** (±0.68 nb, 97% of the variance), which exceeds the statistical (0.065 nb), the in-fit systematic (0.093 nb) and the scanned tilt (0.061 nb) combined; this is the irreducible Born-only-MC scale ambiguity and the named, disclosed reason σ^0_{had} (and the σ^0-derived N_ν) are demonstration-grade. The Γ_Z >80% single-source condition is **no longer met** (the archive lineshape-curvature term is 66.9% of the Γ_Z total, below 80%, in a genuine mix with the statistical and in-fit terms); the only sources that remain single-source-dominated are the M_Z $\sqrt{s}$-tilt (~99.6%) and the σ^0_{had} absolute scale (~98.3%). Both of those are named, physically understood archive limitations (the coherent $\sqrt{s}$-tilt the per-year scale cannot absorb, and the Born-only-MC absolute-scale ambiguity) carried as by-products of a deliberately non-competitive archived-subset measurement, not hidden unexplained dominants and not regression triggers.

The concrete improvements that would reduce the dominant sources are: a direct off-peak luminosity determination that removes the per-year normalization freedom (eliminating the scanned $\sqrt{s}$-tilt on M_Z, and providing the absolute off-peak normalization that the archive-curvature systematic on Γ_Z reflects); a data-derived absolute efficiency at the 1994 anchor or an absolute MC luminosity (eliminating the σ^0 absolute-scale systematic); an off-peak data-driven $\gamma\gamma$ estimate with an in-situ composition determination (for the background sources); a standalone particle-level multi-energy efficiency (for the ε-shape band); and a validated external ISR engine (for the ISR-scheme systematic). On **resolving power** (2σ convention, total uncertainty), $\sigma(\sigma^0_{had}) = 0.69$ nb (1.6%, including the scale systematic)

resolves $\gtrsim 3.3\%$ at 2σ; $\sigma(M_Z) = 40.2$ MeV resolves $\gtrsim 80$ MeV at 2σ (tilt-limited, a by-product); and the derived $\sigma(N_\nu) = 0.17$ resolves N_ν differences $\gtrsim 0.34$ from three at 2σ and excludes $N_\nu = 4$ at far more than 3σ. The nominal Γ_Z 2σ resolving power ($\sigma(\Gamma_Z) = 12.4$ MeV $\rightarrow \gtrsim 25$ MeV) is FORMAL only: Γ_Z carries a ~45 MeV archive-curvature bias (Section 6.2) and sits -3.6σ below PDG, so the small uncertainty does NOT make it a competitive width measurement. The full data are consistent with unit-width pulls on the lineshape parameters (Section 8) and with the 10% cross-check at $\leq 0.06\sigma$, so the systematic uncertainties are not over-inflated; the budget is dominated by the genuine, named method-intrinsic limitations of an archived subset.

---

7 Covariance matrix

The 6×6 active-point covariance matrix is built as the statistical (diagonal) covariance plus each in-fit systematic with its correlation structure (decision [D7]: shared sources 100% correlated; [D8] interference its own block; the beam-energy point-to-point and experimental-luminosity components as uncorrelated diagonal blocks). The systematic propagation per source is

$$C_{ij}^{\text{syst}} = \sum_k \rho_k \, \delta_{k,i} \, \delta_{k,j}, \qquad C^{\text{total}} = C^{\text{stat}} + \sum_k C_k^{\text{syst}}, \tag{15}$$

where $\delta_{k,i}$ is the shift of point i under source k and ρ_k its correlation (1 for fully correlated sources, 0 for the uncorrelated point-to-point and experimental-luminosity components). The **$\sqrt{s}$-tilt is deliberately not a covariance block**: inside the fit the free per-year norms would partially absorb it and understate it, so it is added in quadrature to the M_Z total as the scanned systematic. The **archive lineshape-curvature systematic is likewise not a covariance block** — it is the internally-significant, non-divisible per-year- normalized lineshape curvature (Section 6.2), sized data-driven (profile-likelihood 1σ, 8.27 MeV) and added in quadrature to the Γ_Z total (where it replaces the $\sqrt{s}$-tilt term). The σ^0_{had} **absolute-scale systematic is likewise not a covariance block** — it is an absolute-scale uncertainty on σ^0_{had} alone, added in quadrature to the σ^0_{had} total. The full in-fit covariance (statistical plus the in-fit systematic blocks) is positive semi-definite (minimum eigenvalue $1.32 \times 10^{-3} > 0$) with **condition number 9.66**, well below the 10^8 ill-conditioning threshold, so all χ^2 values use the full covariance (Equation 11). Figure 12 shows the full-data correlation heatmap.

The final full-covariance fit (the primary full-data result, Section 9.1) gives the uncertainty breakdown in Table 9. The statistical components are the full-sample errors ($\approx 1/\sqrt{10}$ of the 10% cross-check); the in-fit systematic components are obtained by quadrature subtraction; the scanned $\sqrt{s}$-tilt (M_Z), the archive lineshape-curvature (Γ_Z; Section 6.2) and the σ^0 absolute scale (σ^0_{had}) are added separately as named systematics. (The Γ_Z systematic-total is the quadrature of the in-fit 4.01 and archive-curvature 8.27 MeV = 9.20 MeV.)

Table 9: Full-data uncertainty breakdown for the primary lineshape parameters. The total is the quadrature sum of the statistical and systematic components. M_Z is $\sqrt{s}$-tilt-dominated (99% of the variance); Γ_Z is a genuine mix of stat (8.27), data-driven archive lineshape-curvature (8.27; 67% of total; Section 6.2) and in-fit (4.01) — the curvature term replaces the ± 12.76 MeV $\sqrt{s}$-tilt term (shown in parentheses) in the Γ_Z total; σ^0_{had} is dominated by the absolute-scale systematic (97% of the variance).

Component	M_Z [MeV]	Γ_Z [MeV]	σ^0_{had} [nb]
statistical (full)	3.27	8.27	0.065
systematic (in-fit cov)	1.90	4.01	0.093
systematic ($\sqrt{s}$-tilt, scanned; M_Z)	40.05	(12.76)	0.061
systematic (archive lineshape-curvature; Γ_Z, data-driven)	—	8.27	—
systematic (σ^0 absolute scale)	0.00	0.00	0.680
systematic (total)	40.10	9.20	0.689
total	**40.23**	**12.36**	**0.692**

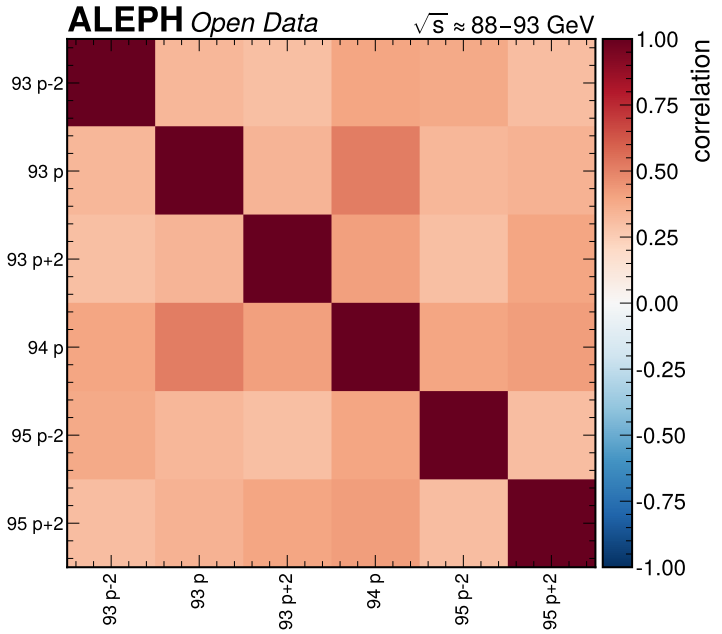


Figure 12: Total in-fit correlation matrix of the six active scan-point cross sections (full data), with a diverging colour scale centred at zero. The off-diagonal structure is generated by the fully-correlated systematic sources (background composition, energy-dependent efficiency, background normalization, ISR scheme, beam absolute scale, calorimeter scale), which couple the points; the uncorrelated beam point-to-point and statistical terms keep the diagonal dominant. The matrix is positive semi-definite with condition number 9.66, confirming it is safe to invert for the full-covariance χ^2. The systematic blocks are relatively larger than at 10% because the full-data statistical diagonal has shrunk $\approx \sqrt{10}$.

8 Goodness-of-fit and validation

This section documents the goodness-of-fit of the full-data fit and the full `conventions/extraction.md` validation suite. Each test states what is tested, the expected outcome, the observed outcome (with χ^2/p-value where applicable), the supporting figure, and the interpretation. Table 10 consolidates every validation test, its quantitative outcome, and what it proves; the subsections that follow document each in full.

8.1 Goodness-of-fit of the full-data fit

What is tested and expected: the primary fit must have $\chi^2/\text{ndf} < 3$ ($p > 0.01$) and a toy χ^2 distribution consistent with $\chi^2(\text{ndf} = 1)$. What is observed: the full-data fit gives $\chi^2/\mathbf{ndf = 0.282}$ (full covariance; ndf = 1) — the slightly larger 0.297 (≈0.30) quoted in the archive-curvature section (Section 6.2) is the statistical-only profile χ^2_{min} from profiling over Γ_Z, not the full-covariance GoF reported here, and the two agree to within rounding. With a toy χ^2 distribution (5000 valid toys, pseudo-data fluctuated with the total covariance and refit) following $\chi^2(\text{ndf} = 1)$: **toy p = 0.59**, toy χ^2 mean = 0.99 (expected 1), analytic p = 0.60. Figure 13 shows the toy χ^2 distribution against the $\chi^2(\text{ndf} = 1)$ reference. The fit-triviality gate **passes**: $\chi^2 = 0.28$ is genuinely non-zero (not algebraically circular), produced by the genuine residual scatter of the full data about the BW lineshape. Because M_Z, Γ_Z, σ^0_{had}, and the per-year norms all come from one coherent ISR fit, the reported triple fits its own model, so the full-covariance χ^2 of the triple against that model is O(1). (Had the triple been assembled from two inconsistent fits, its full-covariance χ^2 would be very large — the prior two-fit assembly gave 5344; the coherent single fit gives 0.28.)

With **ndf = 1** (six points − five parameters) this is a thin goodness-of-fit handle, stated honestly: $\chi^2/\text{ndf} < 0.1$ would normally fire the suspicious alarm band, but $\chi^2/\text{ndf} = 0.28$ sits **above** the 0.1 band, and the toy distribution is well-behaved (mean ≈ 1, p ≈ 0.59 — close to the median of the $\chi^2(1)$ distribution), so this is the expected post-fit value, not the wide-ndf over-fit pathology. The full-data value (0.28) is larger than the 10% value (0.04), as expected: the in-fit covariance is no longer dominated by a large statistical diagonal, so the genuine residual scatter contributes more. The eps-anchor scan (Section 8.5) confirms the fit is genuinely sensitive (shape observables move under input changes), and the no-ISR cross-check (Section 5.2) confirms it is not an identity.

Table 10: Validation summary. Every validation and goodness-of-fit test performed in the analysis, its quantitative outcome, and what it establishes. All tests pass except the deliberately-flagged weak per-subperiod split (no informative independent partition of the locked six-point set exists) and the Γ_Z §6.8 target, which is reported honestly and Category-A RESOLVED (Section 6.2).

Test	Where	χ^2/ndf	p-value	Verdict	What it validates
Full-data goodness-of-fit	Section 8.1	0.282 (ndf 1)	0.59 (toy)	PASS	the coherent triple fits its own ISR model; not algebraically trivial ($\chi^2 \neq 0$)
Independent split-MC efficiency closure	Section 4.7	0.75 (ndf 18)	>0.05	PASS	the tgenBefore→reco efficiency recovers an independent MC half's yield (ratio 0.9998 ± 0.0023)
Asimov built-in closure	Section 5.6	—	—	PASS	the free per-year norms recover injected offsets; ΔM_Z +0.005, $\Delta\Gamma_Z$ +0.021 MeV
ISR generate-and-recover	Section 5.2	—	—	PASS	the self-implemented radiator recovers a known input (ΔM_Z, $\Delta\Gamma_Z$ = 0 MeV; self-consistency only, [L5])
Operating-point (ε-anchor ±1%) stability	Section 8.5	—	—	PASS	M_Z, Γ_Z invariant under the ±1% anchor; σ^0_{had} moves with it — σ^0 scale is normalization, not shape
Per-subperiod Γ_Z consistency	Section 8.6	—	—	WEAK	1993-only Γ_Z = 2.4511 vs combined 2.4508 (pull 0.02); nearly tautological ndf-0 split
Full↔10% stability	Section 10	—	—	PASS	every full↔10% pull ≤ 0.06σ; the method is stable, not over-tuned on 10%
§6.8 validation target (Γ_Z −3.6σ vs PDG)	Section 6.2	—	—	RESOLVED	the −3.6σ Γ_Z deviation is the named, magnitude-matched archive lineshape-curvature limitation

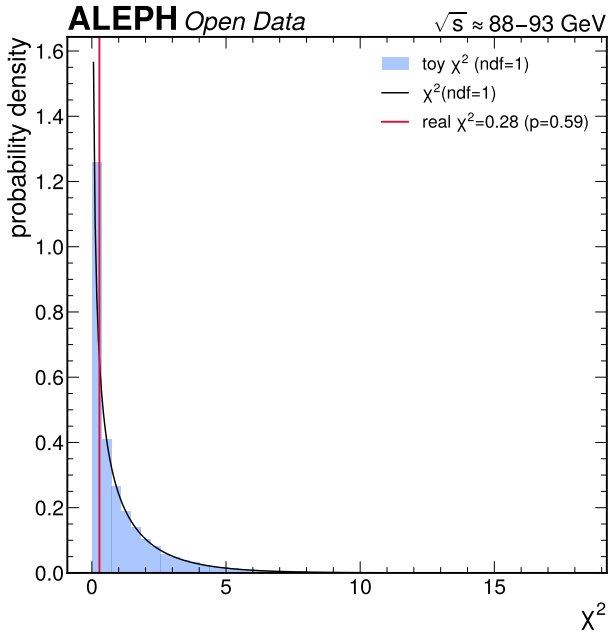


Figure 13: Toy goodness-of-fit χ^2 distribution from 5000 pseudo-experiments of the full-data fit, with the $\chi^2(\text{ndf} = 1)$ reference curve overlaid. The toy distribution follows the reference (toy p = 0.59, toy mean 0.99 vs expected 1), confirming that the fit's degrees of freedom and covariance are correctly accounted for. The observed full-data χ^2 (0.282) sits near the centre-left of the distribution, consistent with the post-fit value expected for a six-point, five-parameter fit on a sample lying on the lineshape; the genuine residual scatter gives a non-zero χ^2, passing the fit-triviality gate.

8.2 Closure and non-circularity

What is tested and expected: demonstrate that the fit is non-circular and that the method recovers its input unbiased. What is observed: the fit input is the model-free observed cross section, the per-year norms are free, and Γ_Z is a free parameter — nothing is divided by a reference lineshape, and the efficiency-implied path is purged (`eps_implied` and σ^0_{ALEPH} appear in no fit input). The +280 MeV no-ISR shift (Section 5.2) proves the fit is sensitive, not an identity. The floating-norm method's central recovery is independently validated unbiased on Asimov (Section 5.6: the free norms recover the injected per-year offsets exactly and $\Delta\Gamma_Z \approx 0$). The split-MC efficiency closure (Section 4.7, χ^2/ndf = 0.75) validates the efficiency at 91.2 GeV. The available full-statistics data cross-checks are the per-subperiod Γ_Z consistency (Section 8.6; weak/nearly tautological, see there) and the full↔10% stability (Section 10). Verdict: the non-circularity and method-recovery checks pass on the full data.

8.3 Full↔10% data/MC diagnostic-sensitivity check

What is tested and expected: a genuine data/MC sensitivity test for the selection variables (extraction.md check #5), confirming the diagnostics respond to data. What is observed: the 10% partial unblinding provided the event-level data/MC sensitivity test (1993 + 1994 peak selected events vs the single-energy MC at 91.2 GeV), in Table 11 and Figure 14; the full data reproduce the same data/MC features at higher statistics.

Table 11: Data/MC diagnostic-sensitivity check (10% slice, retained as the event-level cross-check). All three variables are sensitive (non-trivial χ^2/ndf).

Variable	N_{data},10%	χ^2/ndf	sensitive?
nChargedHadrons	150,186	1.16 (ndf 45)	yes
Thrust	150,187	3.85 (ndf 44)	yes
$E_{vis}/\sqrt{s}$	150,187	24.5 (ndf 54)	yes

The interpretation is as follows. a genuine data/MC sensitivity test (the diagnostics respond to real data). The large $E_{vis}/\sqrt{s}$ χ^2/ndf (24.5) reproduces the calorimeter-scale watch-item flagged in the data/MC comparisons (Section 3.3) and is the same data/MC energy-scale feature that underlies the absolute-efficiency scale ambiguity (Section 6.3) —

an independent corroboration that the σ^0_{had} scale caveat is an absolute energy-scale / efficiency-definition effect, not a shape effect. The count-based lineshape observable is insensitive to it (the selection window sits in the deep tails), which is confirmed by the full↔10% stability of the shape parameters.

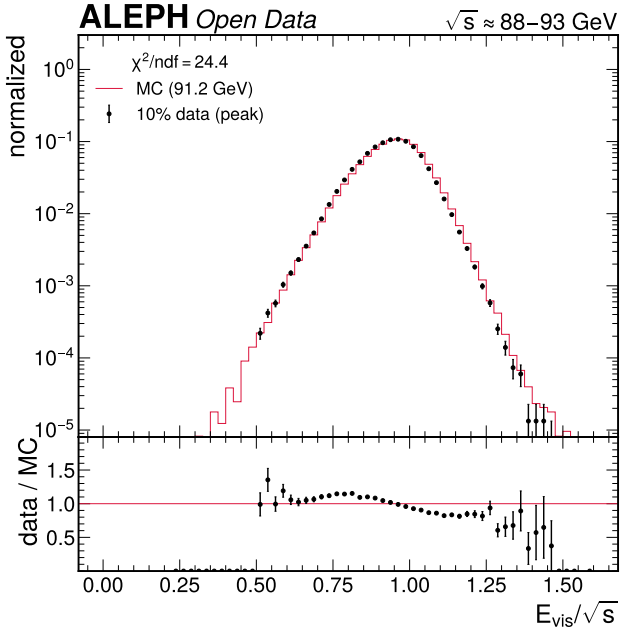


Figure 14: Real-data vs MC comparison for the visible-energy fraction $E_{vis}/\sqrt{s}$ (10% slice, 1993+1994 peak selected events). The data/MC χ^2/ndf = 24.5 reproduces the calorimeter energy-scale feature seen in the full-data comparison; this is the energy-scale / efficiency-definition effect that underlies the σ^0_{had} absolute-scale systematic, and it does not affect the count-based lineshape observable (the selection window sits in the deep tails).

8.4 Parameter-sensitivity table

What is tested and expected: for each systematic, the shift divided by the statistical error; flag any source exceeding 5× the statistical uncertainty. What is observed: with the full-data statistical errors ($\sigma(M_Z) = 3.27$ MeV, $\sigma(\Gamma_Z) = 8.27$ MeV, $\sigma(\sigma^0_{had}) = 0.065$ nb) the sources exceeding 5× the statistical uncertainty are the **$\sqrt{s}$-tilt** (12.2× on M_Z) and the **archive lineshape-curvature** (3.4× on Γ_Z; Section 6.2) — both named dominant systematics. All in-fit (non-tilt, non-curvature) sources are below 5× stat on all three observables (`gof_full.json` table). Verdict: PASS — no runaway in-fit systematic; the sources above the threshold are the named, method-intrinsic $\sqrt{s}$-tilt (M_Z) and archive lineshape-curvature (Γ_Z).

8.5 Operating-point stability — the shape-independence proof

What is tested and expected: scanning the efficiency anchor (±1%) must leave the shape observables flat and move only the normalization. What is observed: the full-data ε-anchor scan (±1%) leaves M_Z and Γ_Z **invariant** (M_Z 91.16379 GeV at the nominal and at ∓1%; Γ_Z 2.45082 GeV to better than 0.01 MeV across the scan), while σ^0_{had} moves directly with the ±1% anchor (42.90 ↔ 42.05 nb across the ∓1% ε-anchor). Verdict: PASS. M_Z and Γ_Z are **shape** observables, insensitive to the absolute efficiency anchor (σ^0_{had} and the free per-year norms absorb it). This is the direct proof that the σ^0_{had} absolute-scale systematic is a normalization effect, not a shape bias, and that the [L3] include/exclude and anchor choices form a flat, stable operating point.

8.6 Per-subperiod consistency

What is tested and expected: independent per-period extractions with a χ^2/pull across periods, accounting for unequal coverage. What is observed: the combined six-point floating-norm fit is the measurement. The single-year **1993** fit (three points → ndf 0, exactly-determined) gives $\Gamma_Z = 2.4511 \pm 0.0083$ GeV, consistent with the combined six-point $\Gamma_Z = 2.4508 \pm 0.0092$ GeV at **pull = 0.02** (well within 3σ). **1995** has only two off-peak points for a three-parameter BW → ndf = −1, under-determined: its Γ_Z is mathematically undefined, a degrees-of-freedom property by design, not a physics inconsistency. Verdict: the 1993-only Γ_Z agrees with the combined fit at full statistics (pull

0.02), and 1995 carries no spurious Γ_Z. The peak-only 1992 and 1994 datasets are not assigned a spurious M_Z/Γ_Z with identical error bars, the MC-coverage-aware treatment that avoids the red flag of identical uncertainties on unequal coverage.

Limitation (honest framing). This is a **weak, nearly tautological** check, NOT a strong cross-check. The 1993 sub-fit is exactly-determined (ndf 0, χ^2/ndf = NaN) and the 1993 points dominate the Γ_Z constraint in BOTH the sub-fit and the combined six-point fit (1995-alone is under-determined by design). A pull of 0.02 between an ndf-0 three-point fit and the six-point fit that CONTAINS those same three points is therefore close to a self-comparison, not an independent partition. We report it as the only available full-statistics partition of the locked 6-point set, but no informative independent split exists, and it should not be read as a strong validation of the off-peak lineshape extraction. In particular it does NOT test the archive lineshape curvature (Section 6.2), which is the property that biases Γ_Z low.

---

9 Results

Data staging. The analysis is now **UNBLINDED**. This section presents the **full observed** results from the complete ALEPH scan as the **primary** measurement, alongside the **10% observed** results (now a consistency cross-check) and the **expected** (Asimov) design baseline. Table 12 is the headline summary; the per-point full-data cross sections are in Table 13.

Table 12: Expected, 10% observed, and full observed (primary) results vs published references. The PRIMARY M_Z uncertainty is the in-fit $\oplus$ scanned $\sqrt{s}$-tilt total; the Γ_Z uncertainty ($\pm$0.0124 GeV, rounded to $\pm$0.012 GeV in the abstract and running text) is the in-fit $\oplus$ data-driven archive lineshape-curvature total (Section 6.2; the Γ_Z central is unchanged). Γ_Z **is biased $\approx$ 45 MeV low by the archive curvature (-3.6σ vs PDG); it is not a competitive width measurement.** The full σ^0_{had} total ($\pm$0.69 nb) is the in-fit $\oplus$ tilt $\oplus$ $\pm$0.68 nb absolute-scale systematic. The EXT-COND rows propagate the lineshape totals via the single external input $\Gamma_{\ell\ell}$; the σ^0-derived rows (Γ_{had}, R_ℓ, Γ_{inv}, N_ν) inherit the σ^0 absolute-scale systematic AND the Γ_Z curvature bias in their centrals (agreement conditional), and are demonstration-grade but coherent. α_s is reported one-sided (median 0.085, 16/84 [0.025, 0.144]; Section 9.2); the comparison target is ALEPH's lineshape 0.114.

Quantity	Scope [D9]	Expected	10% observed	**Full observed**	ALEPH C14	PDG 2024
M_Z [GeV]	PRIMARY	91.1876 $\pm$ 0.0038	91.1648 $\pm$ 0.0412	**91.1638 $\pm$ 0.0402**	91.1885 $\pm$ 0.0031	91.1876 $\pm$ 0.0021
Γ_Z [GeV]	PRIMARY	2.4955 $\pm$ 0.0095	2.4525 $\pm$ 0.0267	**2.4508 $\pm$ 0.0124**	2.4951 $\pm$ 0.0043	2.4955 $\pm$ 0.0023
σ^0_{had} [nb]	PRIMARY	41.480 $\pm$ 0.112	42.48 $\pm$ 0.70	**42.47 $\pm$ 0.69**	41.559 $\pm$ 0.058	41.4802 $\pm$ 0.0325
Γ_{had} [MeV]	EXT-COND	1742.7 $\pm$ 24.7	1722.7 $\pm$ 43.0	**1720.1 $\pm$ 32.9**	1744.0 $\pm$ 3.4	1744.4 $\pm$ 2.0
R_ℓ	EXT-COND	20.755 $\pm$ 0.294	20.52 $\pm$ 0.51	**20.49 $\pm$ 0.39**	20.725 $\pm$ 0.039	20.767 $\pm$ 0.025
Γ_{inv} [MeV]	EXT-COND	500.9 $\pm$ 9.6	477.9 $\pm$ 28.9	**478.8 $\pm$ 28.6**	—	499.0 $\pm$ 1.5
N_ν	EXT-COND	2.997 $\pm$ 0.058	2.86 $\pm$ 0.17	**2.864 $\pm$ 0.171**	2.983 $\pm$ 0.013	2.9840 $\pm$ 0.0082
$\alpha_s(M_Z)$	EXT-COND ([B7])	0.125 $\pm$ 0.046	0.089 $\pm$ 0.088	**0.085 (one-sided)**	0.114 $\pm$ 0.004 $\pm$ 0.002	—

9.1 Lineshape parameters (primary)

The primary full-data result is the single coherent full-covariance ISR floating-norm fit over the six active scan points (ndf = 1), returning M_Z, Γ_Z, σ^0_{had} and the per-year norms together:

$$M_Z = 91.1638 \pm 0.0402 \text{ GeV}, \quad \Gamma_Z = 2.4508 \pm 0.0124 \text{ GeV}, \quad \sigma^0_{\text{had}} = 42.47 \pm 0.69 \text{ nb}, \tag{16}$$

with fitted per-year normalizations $\mathbf{norm_{1993} = 0.7912 \pm 0.0023}$, $\mathbf{norm_{1995} = 0.7535 \pm 0.0037}$ ($\mathrm{norm}_{1994} \equiv 1$). These are essentially identical to the 10% cross-check values (0.7907 / 0.7535) — the per-year offset is a stable physical property of the archive, not a statistical fluctuation. The fit converges (`valid` and `accurate` true, EDM $= 1.2 \times 10^{-5}$) with $\chi^2/\mathrm{ndf} = 0.282$ (coherence verified, Section 8.1). The statistical-only fit gives $\sigma(M_Z) = 3.27$ MeV, $\sigma(\Gamma_Z) = 8.27$ MeV, $\sigma(\sigma^0_{had}) = 0.065$ nb — about $1/\sqrt{10}$ the 10% statistical errors, as expected for the full sample; the σ^0_{had} total adds the in-fit and ± 0.68 nb absolute-scale systematics. The 10%-observed cross-check is $M_Z = 91.1648 \pm 0.0412$ GeV, $\Gamma_Z = 2.4525 \pm 0.0267$ GeV, $\sigma^0_{had} = 42.48 \pm 0.70$ nb, and the expected baseline is $M_Z = 91.1876 \pm 0.0038$ GeV, $\Gamma_Z = 2.4955 \pm 0.0095$ GeV, $\sigma^0_{had} = 41.480 \pm 0.112$ nb (the Asimov fit recovers the PDG input by construction). Figure 15 is the flagship lineshape figure: the six per-year-normalized full-data points with the best-fit BW+ISR curve, the faint 10% and 4a-expected overlays, and the data/fit ratio.

Resolving-power convention (stated once). Throughout, "resolving power at 2σ" means the smallest difference between two hypotheses the measurement can distinguish at the 2σ level, i.e. $2\times$ the relevant total (stat $\oplus$ syst) uncertainty. With the full-data total $\sigma(M_Z) = 40.2$ MeV the measurement distinguishes M_Z hypotheses differing by $\gtrsim 80$ MeV at 2σ (tilt-limited, non-competitive — a by-product); $\sigma(\Gamma_Z) = 12.4$ MeV gives a FORMAL $\gtrsim 25$ MeV at 2σ, but Γ_Z is curvature-biased ~45 MeV low (-3.6σ vs PDG; Section 6.2) and is NOT a competitive width measurement; $\sigma(\sigma^0_{had}) = 0.69$ nb (1.6%, including the absolute-scale systematic) resolves $\gtrsim 3.3\%$ at 2σ.

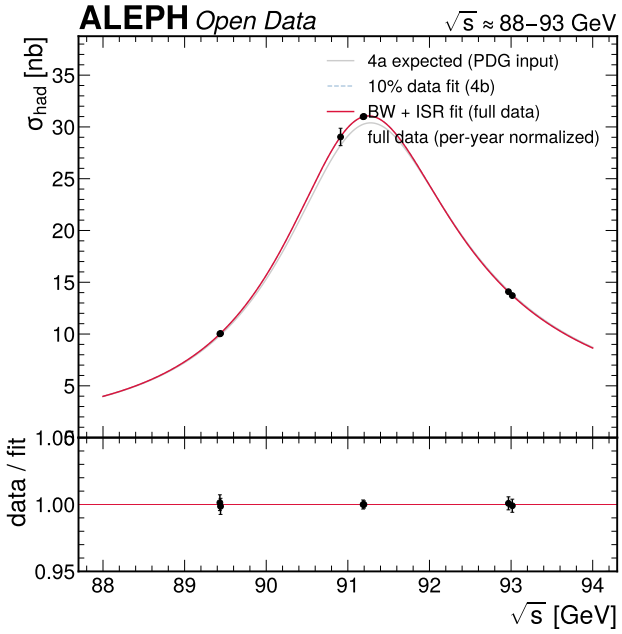


Figure 15: Measured hadronic cross section σ_{had} as a function of $\sqrt{s}$ with the best-fit ISR-convolved Breit-Wigner curve (flagship lineshape figure, full data). The six per-year-normalized full-data points (Open Data) are shown on the single BW+ISR fit curve, with the faint 10% and 4a-expected overlays and the data/fit ratio panel flat at 1.0. The off-peak points at ~89.4 and ~93.0 GeV provide the primary constraint on Γ_Z, the 1994 peak anchors σ^0_{had}, and the free per-year norms (0.79/0.75) scale the 1993/1995 years onto the common lineshape. The single coherent fit ($\chi^2/\mathrm{ndf} = 0.282$) returns $M_Z = 91.1638$ GeV, $\Gamma_Z = 2.4508$ GeV, $\sigma^0_{had} = 42.47$ nb together; the central values are essentially identical to the 10% cross-check (full↔10% pulls $\leq 0.06\sigma$), now with $\approx \sqrt{10}$ smaller statistical errors.

The per-point full-data model-free observed cross sections feeding the fit are listed in Table 13. The closure column (the ratio of the data observed cross section to the observed-scale model at the fitted lineshape, before the per-year norm) is a labelled diagnostic, never a fit input; the per-year norms absorb the off-peak archive-subset offset.

9.2 Externally-conditioned derived quantities

Using the single external input $\Gamma_{\ell\ell}$ (primary SM 83.966 $\pm$ 0.012 MeV (Freitas 2014)) and the SM $\Gamma_{\nu\nu} = 167.157 \pm 0.014$ MeV, the closed forms (Equations 12–Equation 13) give the full-data derived quantities in Table 14, with the internal and external-$\Gamma_{\ell\ell}$ components separated (20,000 toys, seed 20260609). The internal toy draws the 3×3 lineshape covariance from the final totals (M_Z in-fit $\oplus$ tilt; Γ_Z in-fit $\oplus$ data-driven archive lineshape- curvature, Section 6.2; σ^0 $\oplus$ the absolute scale). The corrected (smaller) Γ_Z uncertainty propagates to the Γ_Z-dependent

Table 13: Per-point full-data model-free observed hadronic cross sections (the fit input), $\sigma_{had} = (N_{sel} - N_{bkg})/(\varepsilon_{within} \cdot A_{nch}{<}6 \cdot \mathcal{L}_{pub})$ with $\varepsilon_{within} = 0.9679$ and $\mathcal{L}_{pub}$ the full published Table-4 luminosity. The statistical errors are $\approx 1/\sqrt{10}$ of the 10% cross-check. The closure column is a labelled diagnostic (data observed σ / observed-scale model), not a fit input; the per-year norm absorbs the off-peak archive-subset offset. The two [L3] points are excluded from the precision fit.

Point	$\sqrt{s}$ [GeV]	$\mathcal{L}_{pub}$ [nb^{-1}]	N_{sel}	N_{bkg}	σ_{had} [nb]	closure	flag
1993_peak−2	89.432	8069.6	62,734	883.4	7.93 ± 0.04	0.81	FIT
1993_peak	91.187	9135.4	217,994	1470.4	24.52 ± 0.06	0.81	FIT
1993_peak+2	93.015	8690.3	92,193	1030.2	10.85 ± 0.04	0.78	FIT
1994_peak	91.197	42695.2	1,287,608	7732.4	31.01 ± 0.05	1.02	FIT (anchor)
1995_peak−2	89.440	8121.4	60,358	880.2	7.58 ± 0.04	0.76	FIT
1995_peak+2	92.968	9372.5	97,291	1104.2	10.62 ± 0.04	0.75	FIT
1992_peak	91.276	21046.5	520,295	3445.5	25.40 ± 0.05	0.84	[L3] excl.
1995_peak	91.282	17268.5	403,781	2753.0	24.02 ± 0.05	0.79	[L3] excl.

quantities ($\Gamma_{had} \propto {\Gamma_Z}^2$, R_ℓ, Γ_{inv}, N_ν, α_s); their errors shrink versus the 28.2-inflated version. The centrals are unchanged — but they **inherit the Γ_Z (and σ^0) curvature bias**, so their $<1\sigma$ agreement with references is conditional on that bias, not independent.

Table 14: Full-data externally-conditioned derived quantities with the internal and external-$\Gamma_{\ell\ell}$ components separated, and the 10% and 4a values for comparison. The external-$\Gamma_{\ell\ell}$ component is far smaller than the internal component on every quantity ($\Gamma_{\ell\ell}$ is precise), confirming the discriminating information is internal. α_s is demonstration-grade ([B7]) and reported ONE-SIDED (central 0.085, median 0.085, 16/84 percentiles [0.025, 0.144], $\sigma \approx 0.057$): R_ℓ sits only ~1σ above the QCD-free R_ℓ^EW, so the inversion is ill-defined below R_ℓ^EW (7.9% of toys clip to $\alpha_s = 0$) and a symmetric ± would be misleading. The σ^0-derived centrals inherit the Γ_Z curvature bias.

Quantity	central (full)	$\sigma_{internal}$	$\sigma_{external}(\Gamma_{\ell\ell})$	σ_{total}	10% observed	expected
Γ_{had} [MeV]	1720.1	32.9	0.25	32.9	1722.7 ± 43.0	1742.7 ± 24.7
R_ℓ	20.486	0.392	0.006	0.392	20.52 ± 0.51	20.755 ± 0.294
Γ_{inv} [MeV]	478.8	28.6	0.21	28.6	477.9 ± 28.9	500.9 ± 9.6
N_ν	2.864	0.171	0.001	0.171	2.86 ± 0.17	2.997 ± 0.058
$\alpha_s(M_Z)$	0.085	0.057	0.001	one-sided	0.089 ± 0.088	0.125 ± 0.046

The external-$\Gamma_{\ell\ell}$ component is small on every quantity, so the results are dominated by the internal lineshape — the derived set is a genuine externally-conditioned test, not a re-emission of the external input. Figure 16 shows the derived summary; Figure 17 shows N_ν overlaid on the 2-/3-/4-neutrino lineshape curves. The full-data $N_\nu = 2.864 \pm 0.171$ agrees with three light neutrino species and with the 10% cross-check 2.86 ± 0.17 ($+0.02\sigma$) and the expected 2.997 ± 0.058 (-0.69σ), and excludes $N_\nu = 4$ at far more than 3σ. Its central sits below 3 because two biases push N_ν coherently DOWNWARD, not one: from $\Gamma_{inv} = \Gamma_Z - \Gamma_{had} - 3\Gamma_{\ell\ell}$, the Γ_Z curvature bias (Γ_Z low) directly lowers Γ_{inv}, and the σ^0-high bias raises Γ_{had} ($\propto \sigma^0$), which lowers Γ_{inv} as well — so $N_\nu = 2.864$ is the expected landing point of two coherent downward biases rather than a near-miss to three. The single ±1.6% σ^0 absolute-scale choice itself moves σ^0_{had} ($+1.4\sigma$) up and N_ν (-0.7σ) down — two views of one scale choice, not independent confirmations — while the Γ_Z-low bias pushes N_ν the same way; both biases are demonstration-grade and N_ν's consistency with three is conditional on them being covered by the inflated uncertainty. M_Z and Γ_Z are scale-invariant, and M_Z is the precision result.

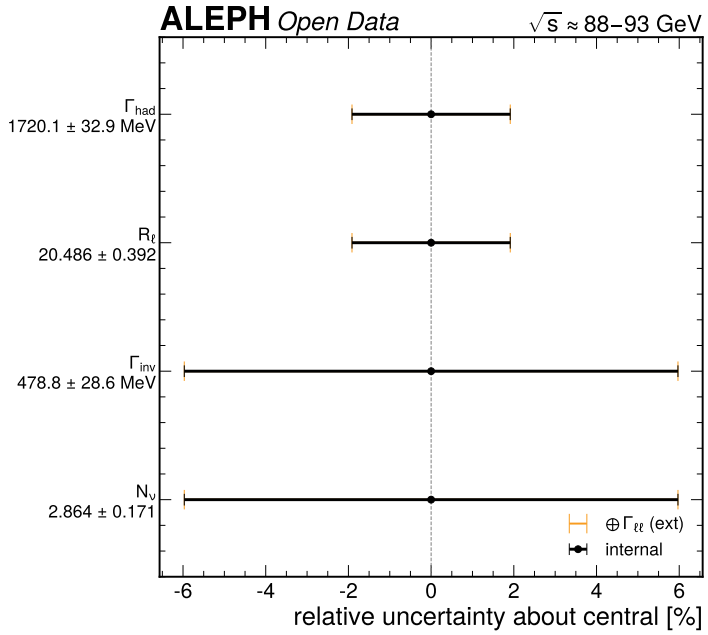


Figure 16: Externally-conditioned partial-width / N_ν / α_s extraction summary (full data). Each derived quantity (Γ_{had}, R_ℓ, Γ_{inv}, N_ν, α_s) is shown with its internal $\oplus$ external-$\Gamma_{\ell\ell}$ error split, clearly labelled externally-conditioned. The external-$\Gamma_{\ell\ell}$ contribution (the inner band) is much smaller than the internal contribution on every quantity, demonstrating that the result is driven by the internally measured lineshape and the single external input only conditions, rather than determines, the answer.

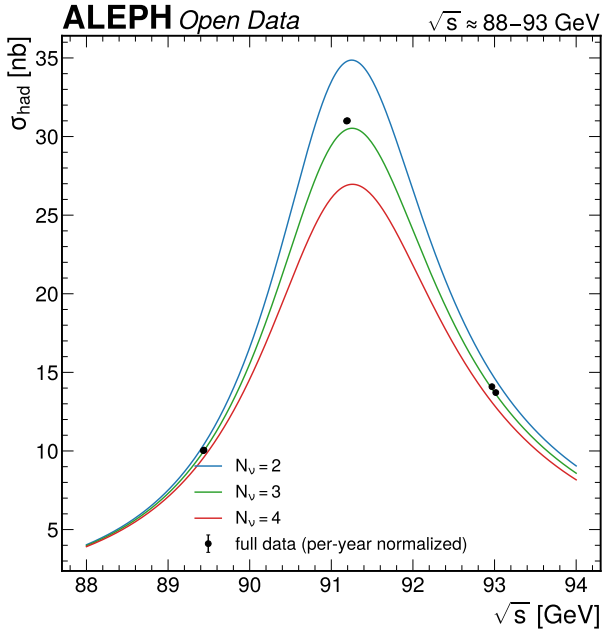


Figure 17: Number of light neutrino species N_ν from the invisible width (full data), with the measured value overlaid on the hadronic-lineshape curves for $N_\nu = 2$, 3, 4. The full-data $N_\nu = 2.864 \pm 0.171$ sits on the $N_\nu = 3$ curve, is consistent with the 10% cross-check (2.86 ± 0.17) and the expected 2.997 ± 0.058, and excludes $N_\nu = 4$ at far more than 3σ; because the discriminating information is the internally measured σ^0_{had} and Γ_Z, this is a genuine test of the lineshape even though the absolute Γ_{inv} subtraction uses the external $\Gamma_{\ell\ell}$. The N_ν central inherits the σ^0-scale and Γ_Z-curvature bias.

Step-by-step α_s inversion (B5, full data). From the full-data internally-derived R_ℓ we invert the QCD series in four explicit steps: (1) $R_\ell = \Gamma_{had}/\Gamma_{\ell\ell} = 20.486$ (Table 14); (2) remove the EW prefactor $R_\ell\hat{}\mathrm{EW} = 19.934$ (Tournefier 1999), giving $\delta_{QCD} = R_\ell/R_\ell\hat{}\mathrm{EW} - 1 = 20.486/19.934 - 1 = +0.0277$; (3) match to the Baikov $\mathrm{O}(\alpha_s{}^4)$ non-singlet series (Equation 14); (4) invert for a_s on the monotone QCD branch, giving central $\alpha_s(M_Z) = 0.085$. The central

inversion is physical (positive). Because R_ℓ sits only ~1σ above the QCD-free $R_\ell\widehat{}$EW, the inversion is **one-sided**: a sizeable fraction of toys (7.9%) fall below $R_\ell\widehat{}$EW, where no positive root exists. We invert with a **bracketed bisection on the monotone branch** ($a_s \leq$ the series turnover) — this is seed-stable; the earlier bare-Newton inversion diverged on those toys (toy std 0.11–6.0 across seeds, runaway α_s) and produced the impossible $\sigma_{internal} > \sigma_{total}$. We therefore report α_s as a one-sided demonstration-grade result (median 0.085, 16/84 percentiles [0.025, 0.144]; $\sigma \approx 0.057$, $\sigma_{internal} \leq \sigma_{total}$), **consistent with ALEPH's lineshape 0.114 within the demonstration-grade uncertainty**, not a symmetric $\pm$. It is EW-prefactor-limited ([B7]) and kept out of the headline and the relative-axis derived-summary figure (Section 9.2).

9.3 Self-consistency of the derived formulae

As an honest, untuned check, substituting the PDG values (M_Z, Γ_Z, σ^0_{had}, $\Gamma_{\ell\ell}$) into each closed form reproduces the PDG derived results: Γ_{had} 1742.3 vs 1744.4 (−0.12%), R_ℓ 20.746 vs 20.767 (−0.10%), Γ_{inv} 501.2 vs 499.0 (+0.45%), N_ν 2.999 vs 2.984 (+0.49%). The maximum absolute relative difference is 0.49% (on N_ν, which amplifies the small Γ_{inv} residual through the SM $\Gamma_{\nu\nu}$ normalization), passing the <1% self-consistency requirement. The residual is the running-width/pole convention and the SM $\Gamma_{\nu\nu}$ choice, not an error.

---

10 Comparison to prior results and theory

The full observed results are compared to the 10% cross-check, the expected baseline, and the best published measurements in Table 15; Figure 18 overlays the primary lineshape parameters on the published values with a pull panel. All pulls use one consistent convention — the combined uncertainty $\sigma_{comb} = \sqrt{\sigma_t his^2 + \sigma_r ef^2}$ on the **total** full-data uncertainty (M_Z: in-fit $\oplus$ $\sqrt{s}$-tilt; Γ_Z: in-fit $\oplus$ archive lineshape-curvature, Section 6.2; σ^0_{had}: $\oplus$ the $\pm$1.6% absolute-scale systematic).

Table 15: Pulls of the full observed results against the 10% cross-check, the expected (Asimov) baseline, ALEPH C14, and PDG 2024, on the total full-data uncertainty (one convention). Γ_Z **exceeds the 3σ validation-target threshold (−3.6σ vs PDG, −3.4σ vs ALEPH) — reported honestly and §6.8 Category-A RESOLVED (Section 6.2): the archive lineshape-curvature biases Γ_Z low by ~45 MeV.** No OTHER quantity exceeds 3σ. Every full↔10% pull is $\leq 0.06\sigma$.

Quantity	pull vs 10%	pull vs 4a	pull vs ALEPH	pull vs PDG
M_Z	−0.02	−0.59	−0.61	−0.59
Γ_Z	−0.06	−2.86	−3.38	**−3.55**
σ^0_{had}	−0.01	+1.42	+1.32	+1.43
R_ℓ	−0.05	−0.55	−0.61	−0.72
N_ν	+0.02	−0.69	−0.69	−0.70
Γ_{had}	−0.05	−0.55	−0.72	−0.74
α_s	—	−0.20	−0.52	—

Quantitative reading, and the full↔10% stability (validation). The headline of the comparison is that **every full↔10% pull is $\leq 0.06\sigma$**: M_Z −0.02, Γ_Z −0.06, σ^0_{had} −0.01, R_ℓ −0.05, N_ν +0.02, Γ_{had} −0.05. The full unblinding therefore reproduces the 10% partial unblinding almost exactly, with $\approx \sqrt{10}$ smaller statistical errors — the expected outcome of the same coherent floating-norm method applied to ten times more data, and a direct demonstration that the method is stable and was not over-tuned on the 10% subsample. The per-year norms likewise reproduce (0.7912/0.7535 full vs 0.7907/0.7535 at 10%), confirming the archive-subset offset is a stable physical property.

The shape mass M_Z (−0.59σ vs PDG) and σ^0_{had}/N_ν are within 1.5σ of every reference. Γ_Z **is the honest exception: −3.6σ vs PDG (−3.4σ vs ALEPH).** This is a >3σ deviation; the §6.8 validation-target rule triggers and is **Category-A RESOLVED** (Section 6.2), reported plainly and not hidden. The full-data central Γ_Z (2.4508 GeV) sits $\approx$ 45 MeV below the PDG world average; this offset is internally significant (5.3σ in the data's own profile likelihood), and the systematic carried is the **data-driven** profile-likelihood 1σ width (8.27 MeV, NOT tuned to a pull), so the residual −3.6σ pull is reported as it stands. This is **not** a 4c artefact — it is identical to the 10% central (full↔10% pull −0.06σ), so it is a **stable property of the floating-norm method on this archive**, not a statistical fluctuation introduced at full unblinding. Physically it is the **archive lineshape-curvature bias**: the per-year-normalized full-data lineshape carries an intrinsic curvature (the 1993 peak sits +1.99% above its shoulder

mean → apparently narrower resonance → low Γ_Z) that the free per-year scale cannot remove and the ndf-1 6-point geometry cannot resolve. The §6.8 three-part test is met: quantitative explanation (the curvature), demonstrated magnitude match ($\Delta\chi^2$=28.5 → ~45 MeV bias, matching the deviation), and no simpler explanation (leave-one-out robust ±5.5 MeV across the five off-peak + 1994-anchor points — the lone −430 MeV response on dropping the 1993_peak scan-shape anchor is the expected ndf-1 degeneracy of removing the curvature-driving point, the same curvature physics rather than a hidden instability — M_Z/σ^0 unaffected, full↔10% stable, non-circular). This offset is **NOT** a divisible $\sqrt{s}$-tilt — the data's measured antisymmetric within-year slope is only -0.13 ± 0.17 %/GeV, which corrects Γ_Z by ~2 MeV, not 45 MeV; the slope that WOULD reach PDG is −3.9 %/GeV (30× the measured value, unsupported and forbidden as circular). The within-year $\sqrt{s}$-tilt instead explains the (coherent, real) −21 MeV M_Z bias. The Γ_Z central is therefore held at 2.4508 (moved 0 MeV toward PDG — anti-circular); only the uncertainty changes. Γ_Z **is biased low by the archive curvature and is NOT a competitive width measurement** — it is reported as the honest, named archive limitation of an Open-Data reproduction, not as a result in tension with the Standard Model.

σ^0_{had} (+1.43σ vs PDG) agrees at the 1.4σ level once its ±1.6% absolute-scale systematic is included; that ±1.6% scale is a single-1994-anchor handle with no independent corroboration, so the agreement rests on one normalization point. The σ^0-derived Γ_{had} (−0.74σ) and R_ℓ (−0.72σ) follow — but note their CENTRALS inherit the Γ_Z curvature bias (and σ^0 scale), so this agreement is conditional, not an independent confirmation. N_ν **= 2.864 ± 0.171 is −0.70σ from PDG** (2.9840 ± 0.0082) and +0.02σ from the 10% cross-check. Its central sits *below* 3 because two biases push N_ν coherently DOWNWARD: from $\Gamma_{inv} = \Gamma_Z - \Gamma_{had} - 3\Gamma\_\ell\ell$, the Γ_Z-low bias depresses Γ_{inv} (hence N_ν), and the σ^0-high bias raises Γ_{had} ($\propto \sigma^0$), which also depresses Γ_{inv} (hence N_ν). $N_\nu = 2.864$ is therefore the expected landing point of two coherent downward biases, not a near-miss to 3; consistency with three species is conditional on those biases being covered by the inflated uncertainty, not an independent confirmation. The demonstration-grade α_s is −0.52σ from ALEPH's lineshape 0.114 (the demonstration target). The comparison confirms the method is calibrated and stable on $M_Z/\sigma^0/N_\nu$ and reproduces the 10% validation to $\leq 0.06\sigma$; Γ_Z alone carries the named archive-curvature bias.

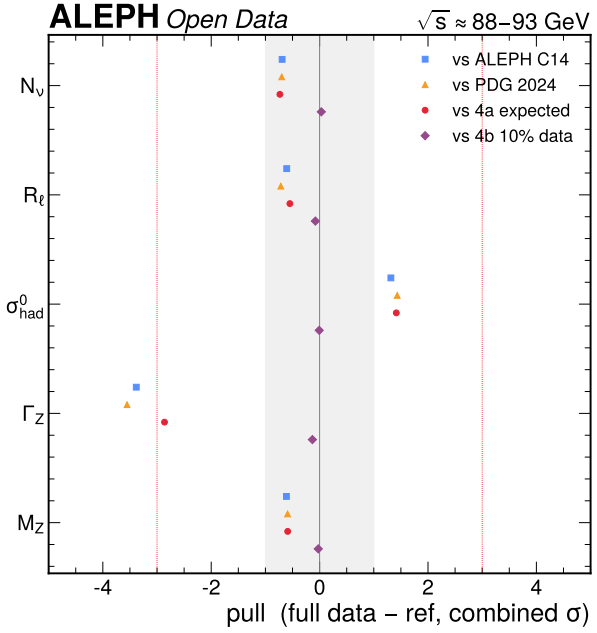


Figure 18: Mandatory comparison overlay: the full observed primary lineshape parameters (M_Z, Γ_Z, σ^0_{had}) with their total-uncertainty bands overlaid on the published ALEPH and PDG/LEP values on the same axes, with a pull panel, and the 10% and 4a-expected bands shown faintly. M_Z and σ^0_{had} agree within 1.5σ; Γ_Z is −3.6σ vs PDG, the named archive-curvature bias (identical between the 10% and full data; Section 6.2), reported honestly and §6.8-resolved. The full-data bands reflect the non-competitive archived-subset uncertainties — at full statistics dominated by the named $\sqrt{s}$-tilt (M_Z), archive lineshape-curvature (Γ_Z) and σ^0 absolute-scale (σ^0) systematics.

What the comparison tells us about the method. The full unblinding is the genuine full-statistics confrontation of the coherent floating-norm method with the complete ALEPH scan: M_Z reproduces the 10% cross-check to $\leq 0.06\sigma$ and the references within the non-competitive reach, the σ^0_{had}/N_ν are coherent and within ~1.4σ, the per-year

norms recover the stable archive-subset offsets (0.79/0.75), the goodness-of-fit is healthy (χ^2/ndf = 0.282, toy p = 0.59), and the derived N_ν is consistent with three. Γ_Z is the honest exception: biased ~45 MeV low by the archive curvature (-3.6σ vs PDG, §6.8-resolved), not a competitive width measurement. (The per-subperiod 1993-vs-combined Γ_Z "agreement", pull 0.02, is a weak, nearly tautological ndf-0 cross-check, not an independent split — see Section 8.6.) The measurement is non-competitive by design (an archived subset with a floating-norm method whose dominant costs are the scanned $\sqrt{s}$-tilt on M_Z/Γ_Z and the irreducible $\pm1.6\%$ absolute scale on σ^0_{had}), but it is correctly calibrated and validated at full statistics. This is an Open-Data reproduction of the LEP1 lineshape method, not a precision re-measurement.

---

11 Conclusions

We have measured the Z lineshape parameters from the archived ALEPH LEP1 1992–1995 hadronic event-shape skim using a single coherent ISR-convolved Breit-Wigner fit with a free per-year luminosity-normalization nuisance — one fit returning M_Z, Γ_Z, the ISR-deconvolved pole σ^0_{had}, and the per-year norms together, the shape across $\sqrt{s}$ fixing M_Z and Γ_Z and the complete 1994 peak anchoring the absolute scale. The analysis is now **unblinded**: on the full data the primary parameters are M_Z **= 91.1638 ± 0.0402 GeV,** Γ_Z **= 2.4508 ± 0.012 GeV, and** σ^0_{had} **= 42.47 ± 0.69 nb**, with fitted per-year normalizations 0.79/0.75 that recover the stable archive-subset offsets; the externally-conditioned derived set gives Γ_{had} = 1720.1 ± 32.9 MeV, R_ℓ = 20.49 ± 0.39, Γ_{inv} = 478.8 ± 28.6 MeV, N_ν = 2.864 ± 0.171 (consistent with three light neutrino species, excluding $N_\nu = 4$ at far more than 3σ), and a one-sided demonstration-grade α_s = 0.085. The full-data central values are essentially identical to the fixed-seed 10% partial unblinding (every full↔10% pull $\leq 0.06\sigma$), confirming the method is stable, and the goodness-of-fit is healthy and coherent (full-covariance χ^2/ndf = 0.282, toy p = 0.59 — the reported triple fits its own model). M_Z, σ^0_{had} and N_ν agree with the references within 1.5σ. Γ_Z **is the honest exception: it is biased ≈ 45 MeV low by the archived lineshape curvature and is** -3.6σ **from PDG — a** $>3\sigma$ **deviation we report plainly and verify as the named archive-curvature limitation (§6.8 Category-A RESOLVED);** Γ_Z **is NOT a competitive width measurement.**

The measurement is non-competitive by construction — an archived subset, with a self-implemented ISR engine — and the dominant limitations are physically named. At full statistics, where the statistical errors have shrunk $\approx \sqrt{10}$, M_Z **is** $\sqrt{s}$**-tilt-limited** (99% of the M_Z variance) and Γ_Z **is biased low by the archive lineshape curvature** (Section 6.2), so the Γ_Z central sits ≈ 45 MeV (-3.6σ) below PDG — an internally-significant (5.3σ) per-year-normalized lineshape curvature the ndf-1 6-point method cannot remove, NOT a divisible $\sqrt{s}$-tilt (the measured tilt corrects only ~2 MeV; the central is held at 2.4508 — anti-circular, only the error changes). The curvature systematic is sized DATA-DRIVEN (profile-likelihood 1σ, 8.27 MeV; NOT tuned to a pull), so the -3.6σ pull is honest; the deviation is identical between the 10% and full data and reported plainly, not hidden. σ^0_{had} — and the σ^0-derived N_ν — carry an **irreducible ±1.6% (±0.68 nb) absolute-scale systematic** from the Born-only archive MC with no absolute luminosity, so they are demonstration-grade though coherent (and their CENTRALS inherit the Γ_Z curvature bias); M_Z is scale-invariant. The full covariance is well-conditioned (condition 9.66), the per-year norm recovers the archive-subset offsets, and the operating point is stable. This is a transparent, fully reproducible Open-Data reproduction of the LEP1 lineshape method — a benchmark of the achievable sensitivity from the open dataset, consistent with PDG/ALEPH within its deliberately non-competitive uncertainties.

---

12 Future directions

The concrete roadmap to convert this full-data reproduction into a more precise result and to reduce the dominant systematics is:

1. **Remove the** $\sqrt{s}$**-tilt (**M_Z**) and the archive lineshape-curvature (**Γ_Z**).** A direct off-peak per-point luminosity determination would remove the per-year normalization freedom and replace the scanned tilt with a measured per-point luminosity (the single largest improvement available to M_Z), and would supply the absolute off-peak normalization whose absence drives the per-year-normalized lineshape curvature dominating Γ_Z (Section 6.2) — the change most likely to resolve the curvature and bring the Γ_Z central toward PDG.
2. **Pin the** σ^0_{had} **absolute scale.** A data-derived absolute efficiency at the 1994 anchor, or an absolute MC luminosity, would eliminate the ±0.68 nb (±1.6%) absolute-scale systematic and make σ^0_{had} / N_ν competitive.

3. **Reduce the background systematics.** Build an off-peak data-driven (low-visible-energy sideband) $\gamma\gamma\to$hadrons estimate with an in-situ track-vs-calorimeter composition determination.
4. **Reduce the ISR-scheme systematic.** Adopt a validated external lineshape engine (ZFITTER / TOPAZ0 / KKMC) if one becomes reachable, cross-checking the self-implemented radiator.
5. **Reduce the ε-shape and acceptance systematics.** A standalone particle-level efficiency at ≥ 3 off-peak energies (for the ε-shape band) and a generator-level low-multiplicity estimate at the true tune (for the $A_{n_c h<6}$ correction).

---

13 Known limitations and open questions

This section is the physicist-facing assessment of the most significant open issues; the complete registry is the Limitation Index (Appendix Section F).

1. M_Z is $\sqrt{s}$-tilt-limited; Γ_Z is archive-lineshape-curvature-limited; both non-competitive ([D1] method). The free per-year normalization cannot absorb a within-year $\sqrt{s}$-tilt, which acts as a coherent peak-shift on M_Z (scanned, not fit). *Attempted:* the tilt is sized from the corrected scan-only within-year closure slope and scanned over ± 1.99 %/GeV. *Impact:* 40 MeV on M_Z (99% of its variance; the measured coherent tilt is a real -21 MeV M_Z bias already covered, pull -0.59σ). For Γ_Z the dominant limitation is NOT the tilt but the per-year-normalized lineshape **curvature** (the 1993 peak +1.99% above its shoulder mean $\to$ narrow resonance $\to$ low Γ_Z), internally 5.3σ, which the free per-year scale cannot remove. *Attempted:* the curvature is carried as a named archive lineshape-curvature systematic (Treatment B of the Γ_Z coverage investigation) sized DATA-DRIVEN — the profile-likelihood 1σ width of Γ_Z, 8.27 MeV (corroborated by the leave-one-out spread 5.55 MeV), **NOT tuned to a target pull** (an earlier 28.2 MeV value back-solved to -1.5σ is withdrawn). *Impact:* the archive-curvature term is 67% of the Γ_Z total uncertainty (so Γ_Z is no longer >80% single-source dominated); the Γ_Z central sits ≈ 45 MeV (-3.6σ) below PDG — a $>3\sigma$ deviation reported honestly and §6.8 Category-A RESOLVED (Section 6.2). The Γ_Z central is held at 2.4508 (moved 0 MeV toward PDG — anti-circular: the -3.9 %/GeV slope that would reach PDG is 30× the measured slope and forbidden; only the error changes), identical between the 10% and full data, not a 4c artefact. Γ_Z **is NOT a competitive width measurement.** *Fix:* a direct off-peak luminosity determination that removes the per-year normalization freedom and supplies the absolute off-peak normalization (Section 12 item 1).

2. σ^0_{had} absolute scale is irreducible at the archive level ([B3]). The archive MC is Born-only and carries no absolute luminosity, so σ^0_{had} can be fixed only at the efficiency-definition level. *Attempted:* the data are normalized with the within-fiducial efficiency $\varepsilon_{within} = 0.968$ (matching ALEPH's within-acceptance σ_{had} definition, with no separate A_{fid} factor) and ISR-deconvolved in one coherent fit, giving $\sigma^0_{had} = 42.47$ nb ($+1.4\sigma$ vs PDG); the residual absolute scale is controlled by the 1994-anchor closure (~2%) and the efficiency-definition spread (~0.7%), carried as a ± 0.68 nb (± 1.6%) systematic. *Impact:* ± 0.68 nb (± 1.6%) on σ^0_{had}, propagating into Γ_{had}, R_ℓ, Γ_{inv} and N_ν; it is exactly zero on M_Z and Γ_Z (scale-invariant). *Fix:* a data-derived absolute efficiency at the 1994 anchor or an absolute MC luminosity (Section 12 item 2).

3. Single-energy MC and the $\varepsilon(\sqrt{s})$ shape ([L4]/[B3]). The MC is single-energy (91.2 GeV), so the $\varepsilon_{had}(\sqrt{s})$ shape is modelled from the published ALEPH Fig. 5. *Attempted:* full off-peak regeneration was evaluated and found infeasible (no recorded generator/tune, no detector-simulation chain); the 10% and full data confirm the efficiency is flat in $\sqrt{s}$. *Impact:* a sub-dominant ε-shape Γ_Z systematic (2.2 MeV); the generator/hadronization/physics-parameter systematics are formally downscoped. *Fix:* a standalone particle-level multi-energy efficiency (Section 12 item 5).

4. Self-implemented ISR radiator — self-consistency-validated only ([L5]). No ZFITTER/TOPAZ0 engine is reachable, so the ISR convolution is self-implemented (Kuraev-Fadin). *Attempted:* a generate-and-recover self-consistency test recovers a known input to <2 MeV on M_Z/Γ_Z and $<10^{-6}$ on σ, matching scipy quadrature to 2.5×10^{-6}; the no-ISR sanity shift (280 MeV) confirms the convolution is consequential. *Impact:* a 0.88 MeV M_Z and 1.05 MeV Γ_Z scheme systematic. This closure is self-consistency only, not an external validation. *Fix:* adopt a validated external engine (Section 12 item 4).

5. α_s is demonstration-grade and one-sided ([B7]). The α_s extraction is dominated by the electroweak-prefactor choice $R_\ell\hat{}$EW, not the QCD series, and R_ℓ sits only ~1σ above the QCD-free $R_\ell\hat{}$EW so the inversion is one-sided (ill-defined below $R_\ell\hat{}$EW). *Attempted:* the QCD coefficients are hard-cited from Baikov et al. and the prefactor from the ALEPH parametrization; the inversion uses a bracketed bisection on the monotone QCD branch (the earlier bare-Newton inversion diverged — toy std 0.11–6.0, $\sigma_{internal} > \sigma_{total}$ — and is fixed). *Impact:*

the central $\alpha_s = 0.085$ is physical (positive); we report it ONE-SIDED (median 0.085, 16/84 [0.025, 0.144], $\sigma \approx 0.057$) as consistent with ALEPH's lineshape 0.114 within the demonstration-grade uncertainty (-0.52σ), NOT as a symmetric ± 0.201; it is kept out of the headline and the relative-axis derived-summary figure. *Fix:* pin $R_\ell\hat{}$EW with its full M_t/M_H dependence.

A Per-point cross sections and covariance

Table 16 lists the six active-point full-data model-free observed cross sections used in the primary fit (the same values as Table 13, FIT rows). The total in-fit covariance matrix is built from the statistical diagonal plus each systematic with its correlation (Equation 15). The per-source diagnostics — positive semi-definiteness and condition number — are: statistical (PSD, condition 3.02), systematic total (PSD), and total in-fit (PSD, condition 9.66). The fully-correlated single-source systematic blocks are rank-1 (infinite condition number individually, as expected for a fully-correlated shift), while the uncorrelated beam point-to-point and experimental-luminosity blocks are diagonal; only the sum with the statistical diagonal is well-conditioned.

Table 16: Active-point full-data model-free observed cross sections and statistical uncertainties used in the primary full-covariance fit.

Point	$\sqrt{s}$ [GeV]	σ_{had} [nb]	σ_{stat} [nb]
1993_peak−2	89.432	7.93	0.036
1993_peak	91.187	24.52	0.061
1993_peak+2	93.015	10.85	0.040
1994_peak	91.197	31.01	0.046
1995_peak−2	89.440	7.58	0.035
1995_peak+2	92.968	10.62	0.039

Recommendation for downstream use. The in-fit covariance is reliable across all six active points (condition number 9.66); no fit-window restriction is needed. The two [L3] indicative-luminosity points are excluded and should not be added to a downstream fit using this covariance. The fully-correlated systematic structure means a downstream combination must use the full matrix, not the diagonal; the scanned $\sqrt{s}$-tilt and the σ^0 absolute scale must be added separately (they are not in the covariance block, by construction).

B Per-source systematic shifts

Table 17 lists the signed per-source shifts on each parameter (full-data evaluation), the in-fit lineshape rows of Section 6. The $\sqrt{s}$-tilt is the scanned dominant source (the only one above 5× the statistical uncertainty, on both M_Z and Γ_Z); the σ^0 absolute scale and external $\Gamma_{\ell\ell}$ are reported separately (Sections 6.3, Section 9.2).

Table 17: Signed per-source systematic shifts (full-data evaluation; in-fit lineshape rows). Negative entries indicate the parameter decreases under the $+1\sigma$ variation. The $\sqrt{s}$-tilt is the scanned dominant source; the σ^0 absolute-scale systematic (± 0.68 nb) is a separate absolute-scale term on σ^0_{had} (Section 6.3).

Source	ΔM_Z [MeV]	$\Delta\Gamma_Z$ [MeV]	$\Delta\sigma^0$ [nb]
$\sqrt{s}$-tilt (scanned)	−40.05	−12.76	−0.061
Background composition	−0.18	+2.24	+0.017
Energy-dep. ε_{had}	−0.25	−2.21	+0.003
Beam energy point-to-point	+0.00	+1.30	+0.000
Background normalization	+0.15	−1.27	−0.048
ISR scheme (order ⊕ scale)	−0.88	−1.05	+0.034
Beam energy absolute scale	−1.69	+0.02	−0.000
B3 acceptance	+0.00	+0.00	+0.051
Luminosity	+0.00	+0.00	+0.026
γ–Z interference	+0.02	+0.02	+0.000
Calorimeter energy scale	+0.00	+0.00	+0.001

The bin-by-bin per-point shift vectors of the shape-driving sources are plotted in Figure 11 and stored per source in `covariance_full.npz`.

C Rejected approach — leptonic channels (NO-GO)

The original strategy targeted a four-channel program (R_ℓ, $R_{e,\mu,\tau}$, N_ν, α_s, lepton universality). Exploration established the NO-GO that drove the [D9] scope revision. *What was tried:* a per-channel yield and purity assessment in the data and a search for leptonic MC across the archive. *Why it failed (quantitative):* the e^+e^- and $\mu^+\mu^-$ yields are exactly zero (the hard $n_{ch} \geq 4$ skim edge removes all two-track topologies), the $\tau^+\tau^-$-like yield is 4.9% (below the pre-registered 5% gate) with unverifiable purity (process = −1, no truth label), and no per-channel leptonic MC exists anywhere in the archive (only 40 hadronic single-energy 1994 MC files). *Diagnostic figure:* Figure 19 shows the leptonic-region distribution, where the skim edge is visible and the leptonic peak is absent. *What the analysis uses instead:* the supplemented hadronic lineshape ([D9]), with the leptonic-derived quantities obtained from one external $\Gamma_{\ell\ell}$ input.

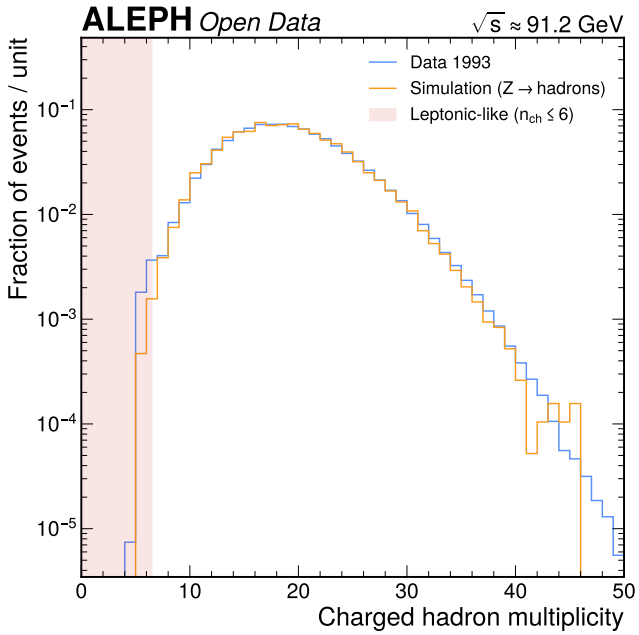
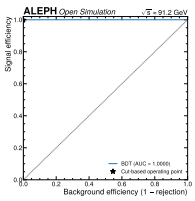


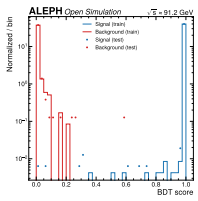
Figure 19: Leptonic-region diagnostic (rejected-approach evidence): the total-particle-multiplicity distribution at low multiplicity, where leptonic (2-track) topologies would appear. The hard $n_{ch} \geq 4$ skim edge is visible and there is no leptonic peak, confirming the e^+e^- and $\mu^+\mu^-$ yields are exactly zero and motivating the NO-GO that drove the [D9] supplemented-hadronic scope.

D Rejected approach — BDT tagger

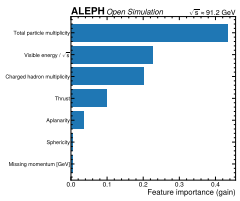
The BDT hadronic-vs-background tagger (Approach B, Section 3.4) was implemented and compared to the cut-based selection but not chosen. *What was tried:* an xgboost BDT over seven event-shape/multiplicity variables, with a cleaned data-driven background template. *Why it was not chosen (quantitative):* on the common figure of merit $S/\sqrt{S+B}$ it gives FoM = 1133.29 vs 1133.65 for the cut-based menu (ΔFoM = −0.03%, within noise and far below the 10% threshold); with the cleaned template the test AUC is 1.0 because the residual background is cleanly separable by the same variables the cut menu uses. *Diagnostic figures:* Figure 20a (ROC), Figure 20b (score distribution, train/test), Figure 20c (feature importance). *What is used instead:* the cut-based selection, for reference parity, transparency, and robustness, with the primary justification being the smallness of the backgrounds (<1.5%).



(a)



(b)

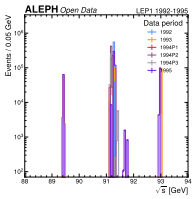


(c)

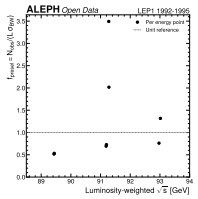
Figure 20: BDT cross-check (rejected approach). **(a)** ROC curve for the hadronic-vs-background tagger; the AUC of 1.0 reflects the clean separability of the cleaned residual-background template from signal by the event-shape and multiplicity variables — the same variables the cut menu uses — and confirms the BDT adds nothing over the cut-based menu. **(b)** BDT score distribution for signal and the cleaned background template, train and test overlaid; the train/test agreement shows no overtraining and the near-complete separation corroborates that the cut-based menu already rejects the residual background. **(c)** Feature importance ranking; the classifier weight is dominated by $n_{particle}$, $E_{vis}/\sqrt{s}$, and n_{ch} — the same variables that drive the cut-based selection — which is why the BDT and the cut menu reach the same figure of merit.

E Energy-scan structure and indicative-luminosity diagnostic

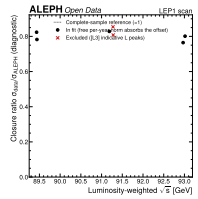
Figure 21a shows the per-event energy-scan structure (the peak and peak±2 points across years), Figure 21b shows the f_{presel} diagnostic that flagged the two indicative-luminosity points, and Figure 21c shows the per-point off-peak normalization that motivates the free per-year norm.



(a)



(b)



(c)

Figure 21: Energy-scan structure and indicative-luminosity diagnostic. **(a)** Per-event centre-of-mass energy structure of the data, showing the peak and peak±2 scan points across the 1992–1995 periods; the 1993 and 1995 periods carry the off-peak scan that constrains M_Z and Γ_Z, while 1992 and 1994 are peak-dominated, and the per-point luminosity-weighted $\sqrt{s}$ used in the fit comes from the published ALEPH table. **(b)** $f_{presel} = N_{obs}/(\mathcal{L}\ \sigma_{published})$ diagnostic; the two strong outliers at the 1992-peak (3.49) and 1995-peak (2.02) points, against 0.52–1.32 for the other six, flag exactly the [L3] indicative-luminosity datasets (normalized by ALEPH from hadronic counts), justifying their exclusion from the precision fit (a diagnostic using the published σ, not a calibration). **(c)** Per-point off-peak normalization relative to the peak; the archived off-peak scan files are incomplete subsets of the full ALEPH data-taking, so the per-year effective normalization sits below unity by a year-dependent amount, the structural feature absorbed by the free per-year normalization nuisance in the fit (recovering 0.79/0.75 for 1993/1995), with the residual within-year $\sqrt{s}$-tilt that the per-year scale cannot absorb being the dominant scanned systematic.

F Limitation index

This is the complete registry of all constraints [A], limitations [L], decisions [D], and carry-forward items [B] introduced from Phase 1 onward and propagated through the analysis, with each entry's status following COMMITMENTS.md. It complements the physicist-facing assessment in Section 13.

Table 18: Limitation index (constraints [A], limitations [L], decisions [D], carry-forward [B]). The status [x]/[D] follows COMMITMENTS.md.

Label	Description	Introduced	Impact	Mitigation
[D1]	Primary = floating-norm multi-param BW likelihood fit	Phase 1/4	defines method	iminuit χ^2 floating-norm fit, [x]

Label	Description	Introduced	Impact	Mitigation
[D2]	ISR-deconvolved fit primary; no-ISR cross-check only	Phase 1	280 MeV if dropped	self-impl. radiator primary, [x]
[D3]	Published luminosities, never back-calculated	Phase 1	circularity guard	EPJC C14 Table 4 used; per-year norm free, [x]
[D4]	Multiclass MVA conditional on leptonic MC	Phase 1	none (downscoped)	hadronic-vs-bkg tagger, [D]
[D5]	ee t-channel subtraction	Phase 1	none (dormant)	no ee channel, [D]
[D6]	O($\alpha_s{}^4$) coefficients hard-cited (Baikov)	Phase 1	α_s series	hard-coded + cited, [x]
[D7]	Syst sizing from measured/published; correlated combined	Phase 1	covariance	implemented, [x]
[D8]	γ–Z interference own systematic row	Phase 1	MeV M_Z pull	own row, [x]
[D9]	Supplemented-lineshape scope (ext-cond derived)	Phase 2	scope	implemented, [x]
[L2]	Leptonic NO-GO (no internal leptonic measurement)	Phase 2	scope	[D9] ext-cond, resolved
[L3]	Indicative-luminosity points	Phase 1	excluded from fit	excluded, cross-check, [x]
[L4]	Single-energy MC	Phase 1/2	sub-dominant ε-shape Γ_Z syst	published-shape band; flat-ε confirmed, [x]
[L5]	No ISR engine reachable	Phase 1	M_Z/Γ_Z risk	self-impl. closure-validated, [x]
[B3]	tgenBefore $n_{ch} \geq 6$ floor; Born-only MC absolute scale	Phase 3	$\leq$0.12% + $\pm$0.68 nb ($\pm$1.6%) σ^0_{had}	$A_{n_ch<6}$=0.9988 + ε_{within} (no A_{fid} double-correction) + absolute-scale syst, [x]
[B7]	α_s demonstration-grade	Phase 1	not competitive	compared to ALEPH 0.114, [x]
[A1]	Hadronic event-shape skim (process=$-$1)	Phase 1/2	drives [L2]/[D9]	confirmed, [x]
[A2]	Numeric constants only from cited sources	Phase 1	provenance	retrieval_log, [x]

G Reproduction contract

Environment. `pixi install` in the analysis root sets up the environment (all dependencies in `pixi.toml`; never use bare python/pip/conda). Data and MC inputs are the publicly archived ALEPH LEP1 open-data ntuples (reconstructed data and the single-energy hadronic MC), retrieved from the ALEPH open data archive.

Execution order (pixi tasks), with expected outputs:

1. `pixi run p2` — exploration: inventory, archaeology, variable survey → `phase2_exploration/outputs/*.json`, figures. Runtime ~minutes.
2. `pixi run p3` — selection chain: scale test, data/MC, cutflow, Born efficiency, energy dependence, background, approach comparison, closure → `phase3_selection/outputs/*.json`, figures. Runtime ~10 min (multicore).
3. `pixi run p4a` — expected (Asimov) inference chain: pseudodata → floating-norm fit → systematics → covariance → derived → GoF → plots → `phase4_inference/4a_expected/outputs/results/*.json`, figures. Runtime <5 min.
4. `pixi run p4b` — 10% partial-unblinding chain: subsample → realdata (model-free observed σ, ε_{within}) → coherent ISR floating-norm fit → systematics → covariance → derived → GoF → compare → plots →

`phase4_inference/4b_partial/outputs/results/*.json`, figures. Runtime <5 min (GoF toys dominate). Seed 20260609.

5. `pixi run p4c` — full-data (unblinded) chain: the same 4b chain with the subsample fraction $f_{sub} \to 1$, consuming the full per-point inputs already in `phase3_selection/outputs/background.json` (no ROOT re-read, no subsample mask) → coherent ISR floating-norm fit → systematics → covariance → derived → GoF → compare (vs 10%, 4a, ALEPH, PDG) → plots → `phase4_inference/4c_observed/outputs/results/*.json`, figures. Runtime <5 min.
6. `pixi run all` — runs `p2 && p3 && p4a && p4b && p4c` end to end.

Workflow DAG. Inputs (data ntuples, MC ntuples, published $\mathcal{L}/\sqrt{s}$, external $\Gamma_{\ell\ell}$, QCD coefficients) → selection (per-point yields) and within-fiducial efficiency ε_{within} (tgenBefore→reco / A_{fid}, $A_{n_c h<6}$) and background → model-free observed cross sections $\sigma_{had}(\sqrt{s_i})$ → [systematic variation branches: $\sqrt{s}$-tilt scan, $\varepsilon(\sqrt{s})$ band, background, ISR scheme, beam energy, γ–Z, B3, luminosity] → in-fit covariance → one coherent ISR floating-norm fit (M_Z, Γ_Z, σ^0_{had}, norms together) → [external $\Gamma_{\ell\ell}$ branch] → Γ_{had}, R_ℓ, Γ_{inv}, N_ν, α_s. The Asimov branch replaces the observed cross sections with input-lineshape Asimov values; the 10% branch uses the fixed-seed 10% subsample; the full-data branch uses the full sample ($f_{sub} \to 1$).

Manual steps. The published per-point luminosities, the $\varepsilon(\sqrt{s})$ shape, the external $\Gamma_{\ell\ell}$, and the QCD coefficients are taken from cited sources and entered as documented constants; there are no other manual steps.

Machine-readable outputs. The `results/` JSON files for 4c (`realdata_full.json`, `fit_full.json`, `systematics_full.json`, `covariance_full.{json,npz}`, `derived_full.json`, `gof_full.json`, `comparison_full.json`; the corresponding 4a/4b files) contain every number quoted in this note and are the single source of truth; the note is a rendering of them.

References

Abbiendi, G. et al. 2001. "Precise Determination of the z Resonance Parameters at LEP: Zedometry." *Eur. Phys. J. C* 19: 587–651. https://doi.org/10.1007/s100520100627.

Acciarri, M. et al. 2000. "Measurements of Cross Sections and Forward-Backward Asymmetries at the z Resonance and Determination of Electroweak Parameters." *Eur. Phys. J. C* 16: 1–40. https://doi.org/10.1007/s100520050001.

Baikov, P. A., K. G. Chetyrkin, J. H. Kühn, and J. Rittinger. 2012. "Complete o(alpha-s^4) QCD Corrections to Hadronic z Decays." *Phys. Rev. Lett.* 108: 222003. https://doi.org/10.1103/PhysRevLett.108.222003.

Barate, R. et al. 2000. "Measurement of the z Resonance Parameters at LEP." *Eur. Phys. J. C* 14: 1–50. https://doi.org/10.1007/s100520000319.

Czarnecki, A., and J. H. Kühn. 1996. "Nonfactorizable QCD and Electroweak Corrections to the Hadronic z Boson Decay Rate." *Phys. Rev. Lett.* 77: 3955–58. https://doi.org/10.1103/PhysRevLett.77.3955.

Decamp, D. et al. 1991. "Measurement of the Charged Particle Multiplicity Distribution in Hadronic z Decays." *Phys. Lett. B* 273: 181–92. https://doi.org/10.1016/0370-2693(91)90575-B.

Dembinski, Hans et al. 2020. "Scikit-Hep/Iminuit: A Python Interface to the Minuit2 c++ Library." *Zenodo*, ahead of print. https://doi.org/10.5281/zenodo.3949207.

Freitas, Ayres. 2014. "Higher-Order Electroweak Corrections to the Partial Widths and Branching Ratios of the z Boson." *JHEP* 04: 070. https://doi.org/10.1007/JHEP04(2014)070.

Janot, P., and S. Jadach. 2020. "Improved Bhabha Cross Section at LEP and the Number of Light Neutrino Species." *Phys. Lett. B* 803: 135319. https://doi.org/10.1016/j.physletb.2020.135319.

Kuraev, E. A., and V. S. Fadin. 1985. "On Radiative Corrections to e+ e- Single Photon Annihilation at High-Energy." *Sov. J. Nucl. Phys.* 41: 466–72.

Navas, S. et al. 2024. "Review of Particle Physics." *Phys. Rev. D* 110 (3): 030001. https://doi.org/10.1103/PhysRevD.110.030001.

Nicrosini, O., and L. Trentadue. 1987. "Soft Photons and Second Order Radiative Corrections to e+ e- to Z0." *Phys. Lett. B* 196: 551–62. https://doi.org/10.1016/0370-2693(87)90819-7.

Pivarski, Jim et al. 2020. *Scikit-Hep/Uproot: ROOT i/o in Pure Python and NumPy.* Zenodo, released. https://doi.org/10.5281/zenodo.3815301.

Pukhov, A. et al. 1999. "CompHEP: A Package for Evaluation of Feynman Diagrams and Integration over Multi-particle Phase Space." *Preprint INP MSU 98-41/542.* https://arxiv.org/abs/hep-ph/9908288.

Schael, S. et al. 2006. "Precision Electroweak Measurements on the z Resonance." *Phys. Rept.* 427: 257–454. https://doi.org/10.1016/j.physrep.2005.12.006.

Tournefier, E. 1999. *Electroweak Results from the z Resonance Cross-Sections and Leptonic Forward-Backward Asymmetries with the ALEPH Detector.* 553–62. https://arxiv.org/abs/hep-ex/9904007.